Spatial Orientation in Virtual Environments

Published and distributed by DUP Science

DUP Science is an imprint of
Delft University Press
P.O. Box 98
2600 MG Delft
The Netherlands
Telephone: +31 15 278 5121
Telefax: +31 15 278 1661
E-mail: dup@library.tudelft.nl

ISBN 90-407-2152-1

Printed in the Netherlands

Spatial Orientation in Virtual Environments

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.F.Wakker,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen
op dinsdag 13 februari 2001 om 16.00 uur

door

Nicolaas Hylke BAKKER

werktuigkundig ingenieur
geboren te Dormaa Ahenkro in Ghana

Dit proefschrift is goedgekeurd door de promotor:
Prof dr.ir. H.G. Stassen.

Toegevoegd promotor:
Dr.ir. P.A.Wieringa.

Samenstelling promotiecommissie:

Rector Magnificus,	Technische Universiteit Delft, voorzitter
Prof.dr.ir. H.G. Stassen,	Technische Universiteit Delft, promotor
Dr.ir. P.A. Wieringa,	Technische Universiteit Delft, toegevoegd promotor
Prof. N.Moray	University of Surrey, Engeland
Prof.dr.ir. R.L. Lagendijk	Technische Universiteit Delft
Prof.dr.ir. F.W. Jansen	Technische Universiteit Delft
Prof.dr.ir. C.A. Grimbergen	Technische Universiteit Delft
Dr.ir. P.O. Passenier	TNO Technische Menskunde, adviseur

Dr.ir. P.O. Passenier en dr.ir. P.J. Werkhoven hebben als begeleiders in belangrijke mate bijgedragen aan de totstandkoming van dit proefschrift.

Dit proefschrift is mogelijk gemaakt door een financiële bijdrage van het instituut voor Toegepast Natuurwetenschappelijk Onderzoek (TNO). Het onderzoek dat wordt gepresenteerd in dit proefschrift werd uitgevoerd bij TNO Technische Menskunde, Soesterberg.

Contents

Acknowledgements

I could not have done this research without the help of many who I would like to thank.

Firstly, I would like to thank my two mentors at TNO Human Factors, Peter Passenier and Peter Werkhoven. Looking in the dictionary for the meaning of the word *mentor* it says: teacher, guide, coach, and counsellor. All these meanings apply to them, and I could not have wished for a better duo.

I thank Henk Stassen my *promotor* and Peter Wieringa my *copromotor* for giving me the trust and the freedom to pursue this research and for their valuable comments on the manuscript.

I would like to thank my colleagues for providing a stimulating work environment in which all the doors are always open and in which hard work goes together with relaxing volleyball, fitness, and bridge sessions. I especially like to thank my room-mates, Patrick Punte and Sjoerd de Vries for their enthusiastic discussions. I thank Hendrik-Jan van Veen for his review of the draft.

I would also like to thank many colleagues who helped with the practical sides of experimenting: Jan van de Kooi for always finding the time to reconfigure and ‘unmount’ the VE lab to make my experiments possible, and for finding many bugs; Wytze Hoekstra for his valuable help at some crucial moments when the programming got rough, and for creating the virtual forest in which many participants got lost; Walter van Dijk, for filming the experiments; Evert Agterhuis for creating just the right amount of noise; Obbe Pranger, Jan Kroonenburg, Lex Viersen, and Jan Benschop for making sure that the apparatus did not fall apart.

Finally, my thanks to the management of TNO Human Factors for giving me the opportunity to do this research.

To my mother and stepfather,
and to my sister and her four guys.

Summary

Spatial Orientation in Virtual Environments

Over the last few decades, human operators who are responsible for the supervision of complex systems are faced with increasing amounts of information. The current graphic interfaces that are used to present information to the operator usually consist of a limited number of two-dimensional computer screens. Navigating within these interfaces places considerable demands on the operator.

Recently, a growing interest can be detected in the application of Virtual Environment (VE) technology as an operator interface. VEs are three-dimensional computer-generated images that can be shown on a conventional monitor, on a large screen display, or on a head-mounted display. Using a VE as an interface provides one with the opportunity to show data in its natural format, for instance the 3D position of an aircraft for air-traffic control, and it gives the interface designer more freedom to arrange and organise data. However, also in these three-dimensional interfaces, the task of finding and retrieving information from the interface may impose considerable task demands on the operator.

Different types of VE technology are available for navigating in these VEs, and different types of navigation can be enabled. A choice has to be made between the two different types of VE interfaces that are available:

- An *immersive* interface that provides rich sensory feedback to the user when moving around in the VE.
- A *non-immersive* interface that provides only visual feedback to the user when moving around in the VE.

When considering the type of navigation, a choice has to be made between two types of displacement:

- *Continuous* displacement in which the viewpoint is moved fluently through the VE.
- *Discontinuous* displacement in which the viewpoint can be moved instantaneously over arbitrarily large distances.

There is insufficient understanding on how these choices may affect the performance of an operator using the interface.

To provide insight into the possible effects of these choices, a qualitative model of human spatial orientation behaviour in a VE was formulated: the Framework for the Investigation of Navigation and Disorientation (FIND). The outline of this model is as follows. In order

to find information, an operator has to determine a movement that has to be executed. To make this decision the operator needs to have knowledge about the environment, and this knowledge is stored in the *cognitive map*. Furthermore, the operator needs to know the *current location*. The current location is determined by a combination of three spatial updating processes: *path integration* that registers the displacement path through the environment on the basis of the available sensory feedback; *visual recognition* that uses the available knowledge in the cognitive map to recognise the location directly from the visual image; *cognitive anticipation* that uses knowledge about the environment and about the intended actions to determine the location after, for instance, a discontinuous displacement. All these processes are controlled by a single *cognitive control process* that distributes attention between the other processes depending on the task.

On the basis of a literature survey structured according to the components of FIND, three main research questions were formulated:

1. When compared to non-immersive navigation, does immersive navigation improve the quality of path integration?
2. When compared to non-immersive navigation, does immersive navigation improve the acquisition of a cognitive map?
3. Does discontinuous displacement affect spatial updating?

To answer these questions a series of nine experiments was carried out to investigate the efficiency of spatial updating and the acquisition of a cognitive map under the influence of different navigation interfaces and different types of displacement.

The results indicate that immersive navigation does indeed improve the quality of path integration. The improvement does not only have an effect when path integration is isolated from other modes of spatial updating, but it also affects spatial updating in the case that visual recognition and cognitive anticipation are also possible. A spatial layout is learned most quickly with an immersive interface. However, as soon as an accurate CM is acquired of the environment, no differences are to be found between the two types of navigation interface.

Discontinuous displacement disrupts spatial updating, leading to an increase in the time needed to acquire a cognitive map. The disorienting effects of a discontinuous displacement can be compensated for by enabling cognitive anticipation of the destination of the displacement to take place. However, some performance decrement remains when this is compared to continuous navigation. The type of discontinuous displacement has an effect on the efficiency of cognitive anticipation.

Supporting good spatial orientation is a prerequisite for the application of VE technology as an interface to support the supervision of complex processes. When deciding on which interface technology should be used, the advantages that were found of immersive navigation have to be considered as well as the disadvantages.

Current immersive VE technology causes eye-strain, headaches and even nausea to many users. These problems have to be solved if widespread use of the technology is to be

allowed. Furthermore, immersive technology is still considerably more expensive than non-immersive technology.
The results of the experiments suggest that immersive navigation might only be beneficial for application domains in which new spatial layouts have to be learned every time or in domains where the primary users are novices. For instance, in training firemen to teach them the layout of new buildings with VE, or in using architectural walkthroughs in VE to show new building designs to potential buyers. For supervisory control applications, immersive navigation will only have an advantage during familiarisation with the interface when the layout of information has to be learned. After knowledge of the layout is acquired, no continuing benefit of immersive navigation should be expected.
When looking at the type op displacement, discontinuous movement should not be allowed when exploring a new environment, because this will hinder the acquisition of a CM. Once the environment is known this recommendation will change. If time is not a critical factor, continuous movement should be clearly preferred. If fast displacement is essential then discontinuous displacement should be preferred. The disorienting effects of discontinuous displacement can be greatly reduced by allowing for cognitive anticipation. The interface designer must make sure that information is provided about the destination of a discontinuous displacement. The type of discontinuous displacements has an effect on the time needed for anticipation. Discontinuous displacements that involve a rotation take more time to anticipate and should, if possible, be avoided.

Recommendations are made for future research and for continuing the investigation of the effects of different types of discontinuous displacement, so that more complete guidelines for the design of VE interfaces can be provided.

Niels Bakker January 2001

1 Introduction

1.1 Background

Over the last few decades, human operators have been confronted with an increasing amount of information, to support the supervision of large-scale, dynamic, and complex systems. Those systems can be chemical factories, ships, telephone networks, or even battlefields that need to be managed. This increase in information is primarily due both to the scale enlargement of the systems and to the advances in information, communication and sensor technology.

The operator in question has to ensure that both the economic and the safety goals of the system are met. Although nowadays most systems are controlled primarily automatically, they still rely on human supervisors to deal with unfamiliar, unanticipated, or abnormal events. To achieve the system goals, the operator needs to combine his perception of the current state of the system with knowledge about the system's properties and knowledge of any possible disturbances so that appropriate action can be determined (Stassen, Johannsen, & Moray, 1990). An important task for the operator in supervisory control applications is to monitor the ever-changing state of the process and diagnose faults that occur (Sheridan, 1988).

Computers are generally used as an interface between the human operator and the system to be supervised. The computer receives data from the process and commands from the supervisor. On the basis of this data, the computer performs automated control actions and presents information to the supervisor. The part of the interface that is used to present information usually consists of a limited number of conventional, two-dimensional, computer screens, displaying graphic representations of both the system properties and the system's state. Several visual representations are required because the amount of information is vast and the display surface limited.

The operator has to gather and to integrate information that is distributed across the information space as presented by the interface. The operator must be able to remember where information is located and he/she must be able to navigate to the required locations, sometimes switching between different representations. The costs of extracting information, in terms of time or the cognitive effort made, must not create an excessive workload for the operator.

Research has shown that navigating through a network of system representations may place substantial demands on the human operator (Schryver, 1994; Billingsley, 1982; Vicente Hayes, & Williges, 1987). The number of representations can be so large that information becomes hard to find (Elm & Woods, 1985). Switching between representations may be confusing because of its discontinuous nature. This may reduce the ability to integrate information appropriately (Woods, 1984). An operator might become fixated on information that is currently displayed and might forget to search for other information, or may misinterpret information because part of the context is missing.

Recently, a growing interest can be seen in the application of Virtual Environment (VE) technology as an alternative interface when it comes to supporting operators, for instance in battlefield management (Dennehy, Nesbitt, & Sumey, 1994), air-traffic control (Wickens & May 1994), and information retrieval from databases (Roth, Chuah, Kerpedjiev, Kolojejchick, & Lucas 1997; Rennison & Strausfeld 1995).

VEs are synthetic sensory experiences that are generated by a computer system with the objective of approximating several attributes of the real world (Kalawsky, 1993). These experiences are usually visual, but may also be auditory, tactile, proprioceptive, or even olfactory. VE systems consist of various interface components that enable natural interaction with the synthetic environment. For instance, if one wants to enable a visual experience to occur, three-dimensional computer-generated images can be shown on a large projection screen surrounding the operator. The images can also be shown on a Head Mounted Display (HMD) in which two small displays are placed just in front of a person's eyes. Movements of the virtual viewpoint in the VE can be slaved to the real head movements by using a head-tracking sensor. Gloves are used, which are equipped with sensors that register the position of the hand and fingers to enable the natural manipulation of virtual objects (Werkhoven & Groen, 1998).

The main application of VE has traditionally been that of simulating real-world environments for the purpose of training, human factors research, product prototyping, and entertainment (Boman, 1995, survey). However, when applying VE as an interface for operators, mimicking reality is no longer the main objective. Instead, the VE is used to access and retrieve information that is arranged in a three-dimensional information-space.

One advantage of VE might be that information that is inherently three-dimensional can be shown in its natural format like, for instance, the 3D position of an aircraft, to support air-traffic control tasks. Alternatively, the extra dimension can be used to encode a specific variable like, for instance, time. Depth can also be used to separate several layers of information. The interface designer is provided with an additional degree of freedom to organise data. The spatial visualisation of data may be a potent aid to human cognitive processing, as the user directly perceives relations in the data (Risch, May, Thomas, & Dowson 1996), rather than having to deduce such relations.

VE potentially reduces the need for the discontinuous switching required with traditional operator interfaces. Instead of switching discontinuously between representations, the operator moves his viewpoint around continuously in a natural way. The natural movement in a VE-interface might reduce the cognitive effort involved in establishing one's own location and finding information. However, efficient displacements over larger distances will probably still necessitate some discontinuous switching of position, to save time.

The use of VE technology has shown that humans still can, in some cases, have considerable difficulty with navigation and spatial orientation compared to moving around in the real world (Ellis, 1993; Kaur, Sutcliffe, & Maiden, 1999). Difficulty with navigating may lead to bumping into objects or even to the inability to reach a destination. People have been reported to loose their general sense of direction, that is to say, to not know where they are (Wilson, Foreman, & Tlauka, 1997). Part of these usability problems may be attributed to the poverty of the visual information that is available in a VE in comparison with the real world, in terms of detail, texture, and resolution.

Besides depending on visual fidelity, it has been suggested that good spatial orientation may also depend on the technology that is used to navigate in the VE (Witmer, Bailey, Knerr & Parsons, 1996; Templeman, Denbrook, & Sibert, 1999; Iwata & Yoshida 1999). A major distinction must be made between immersive and non-immersive VE technology.

Immersive technology enables a high degree of sensory involvement in the virtual environment, allowing the user a sense of presence in the VE. It is possible to distinguish different components of the VE interface each of which may support different kinds of immersion, like the visual display, the navigation interface, or the manipulation interface. An immersive display is present wherever the observer looks, which means that the real world can no longer be seen. The display can be a large screen display completely surrounding the operator or can be an HMD. An immersive navigation interface uses head-trackers to slave the movement of the virtual viewpoint in the VE to the user's head movements. An immersive manipulation interface uses gloves equipped with sensors to make possible the natural manipulation of virtual objects. Non-immersive technology typically consists of a desktop monitor combined with some hand-held input-device, like for instance a joystick or a mouse, to control navigation and manipulation of the virtual objects.

With immersive navigation technology, the operator controls his displacement by moving his head as he would in the real world. Natural sensory feedback from the body is present which can potentially be used for the perception of self-movement (Stassen & Smets, 1995). However, it is not clear how this enhanced movement perception interacts with visual recognition or even cognitive anticipation, which can also be used to determine a location. Attempts to show the benefits of immersion for spatial orientation in VE have often shown no effect (e.g. Waller, Hunt, & Knapp, 1998). With immersion, the supposed increased awareness of one's own location might even turn out to be a disadvantage when a discontinuous displacement is made. Discontinuous displacement creates a discrepancy between actual displacement and perceived movement. No displacement is registered while the location and the visual surrounding is changed. For non-immersive VEs, at least, there is some evidence that discontinuous displacement may lead to spatial disorientation or that it will increase cognitive effort (Bowman, Koller, & Hodges, 1997). However, no literature exists on this possible interaction between the level of immersion and the movement type.

The choice between immersive and non-immersive VE technology has substantial economic consequences. Apart from the graphics renderer, immersion is the main distinction between high-end VE systems on the one hand, using head-trackers combined with HMDs or large projection screens, and low-end VE systems, on the other hand, using conventional desktop monitors with mouse and keyboard. The kind of VE technology that is used to provide operators with an interface to support the supervision of a process needs to enable good spatial orientation and efficient navigation.

To summarise, there is insufficient knowledge about the potential advantages and drawbacks of using immersive or non-immersive VE technology as an interface to support human operators. In order to choose between immersive and non-immersive technology in the implementation of future VE-interfaces, we need to know how such a choice affects the operator's ability to spatially orient himself. In the evaluation of alternative technologies, both the continuous and the discontinuous displacement tasks needs to be considered.

In answer to these questions, a series of experiments is reported which investigates the effects of immersion as opposed to non-immersion on an operator's spatial orientation in both the case of continuous and discontinuous navigation tasks.

1.2 State of the Art

In this section an overview of the literature will be given to show the possible effects that different VE-interfaces may have on operator tasks such as finding and gathering information. In order to understand how the interface can influence task performance, a human behavioural model is desirable. Starting with an existing general model of human information processing, a more specific model will be formulated for the tasks of navigating and finding information. This model will be called the *Framework for the Investigation of Navigation and Disorientation* (FIND). FIND serves three purposes. Firstly, FIND helps to set the boundaries and to define the focus of this research. Secondly, FIND helps to structure the existing literature and reveal unanswered questions. Thirdly, FIND serves to define the manipulations that will be executed during the experiments reported in this thesis.

A literature overview will be given that is structured according to the components of FIND. In this overview, the effects of interface technology on the different functions of FIND will be discussed. On the basis of this section, research questions and expectations will be formulated in the following section, which have to do with the effects of the choice of interface technology on operator task performance.

1.2.1 FIND, a model of navigation behaviour in VE

Numerous psychological models, relating to how humans interact with their environment have been proposed. Wickens (1992) provides a general qualitative model indicating how responses are generated which are based on stimuli from the environment (Fig. 1.1). This model will be used to formulate a more specific qualitative model for the human information processing involved in spatial orientation, but first, a brief description will be provided of the Wickens model.

The first stage in the processing of environmental stimuli is *sensory processing*. The information produced during this stage depend on the characteristics of the sensory receptors that are involved. A representation of the physical stimulus is temporarily preserved in a *short-term sensory store.* An important characteristic of *sensory processing* is that it is automatic, which means that no conscious attention is required for the processing. The next stage of information processing is *perception* or *perceptual encoding*. Going on previous experience stored in memory, information is detected, selected, categorised, or recognised. How this takes place depends on the context or the specific task at hand. The complexity of perceptual encoding may range from single stimulus detection to the more complex recognition of a pattern of features derived from different sensory channels. After perceptual encoding, the *decision and response selection* will determine an appropriate response. This process may vary between being automatic or being extremely complex and requiring careful thought. After a decision has been made, the *response execution* process will determine the necessary muscle commands.

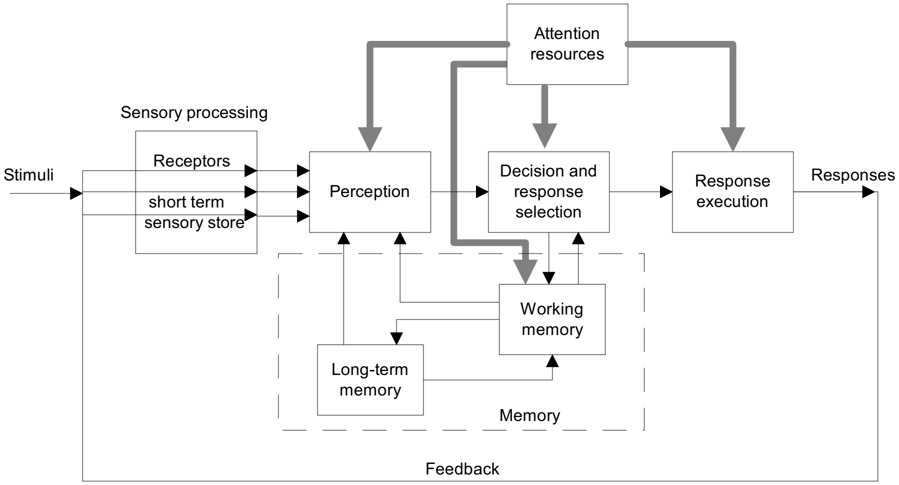


Figure 1.1: A general model of human information processing (adapted with permission from Wickens, 1992); Stimuli from the outside world together with internal feedback stimuli are processed pre-attentively by *sensory processing*. Going on previous experience stored in memory, information is detected, selected, categorised, or recognised. After perceptual encoding, *decision and response selection* has to determine an appropriate response that is subsequently executed. All these processes require *attention resources,* except the sensory processing.

All these processes except the sensory processing require *attention resources*. According to Gaillard (1996), limited resources, such as attention or short term memory, are allocated by a *cognitive control system.* On the basis of performance feedback, the cognitive control system determines which processes need to be activated to achieve the goals. Attention is not only focused *top-down* by the cognitive control system, but can also be attracted *bottom-up* by the salient cues in the environment, for instance a flash of light (Theeuwes, 1992).

Vicente and Rasmussen (1992), discern two levels of cognitive processing. Firstly, analytical processing in which the problem solving is based on symbolic representations, which is serial in nature, requires deliberate attention, and is slow and laborious. Secondly, there is perception-action processing that is parallel, requires little attention, and is fast and effortless. Depending on the amount of experience an operator has, behaviour may shift from being analytical to being perception-action processing oriented.

The mental effort and the amount of attention involved in performing a task depend not only on the task complexity but also on the efficiency of the mental processes that are involved. Mental workload is determined by the proportion of the needed processing capacity in relation to the available processing capacity. Workload is the result of task demands on the one hand and individual factors, like skill level, motivation and emotion on the other hand (Gaillard, 1996).

Let us now see how this general model can be adapted to the more specific case of navigation in a VE interface. Important tasks for the human operator are monitoring, diagnosis and fault-management (Sheridan, 1988). Monitoring means that the operator has to keep track of the ever-changing system state. The operator needs to have an overview of the system state to ensure that the system is functioning normally. In diagnosis and fault management the operator's thinking process frequently switches between different levels of abstraction, changing his need for information (Rasmussen, 1986). Rasmussen distinguishes two main search strategies: symptomatic search and topographic search. With symptomatic search, the operator tries to identify the system state by gathering specific information to confirm or reject a hypothesis about the system's state. With topographic search, the operator focuses on finding the location of a change. Sometimes the operator's tasks may be best supported by local continuous navigation, and sometimes discontinuous navigation might be needed to efficiently gather information from separate regions.

Because this thesis focuses on information retrieval and navigation performed by operators, the *task* is to retrieve a specific item of information. Since it is assumed that the information is spatially organised, information retrieval involves moving to the information location. Therefore, *Response execution* in Wickens' model becomes *movement execution* in FIND (Fig. 1.2). This may involve the control of an input-device. However, before a movement can be executed, a decision has to be made about where to move to. Therefore, *decision and response selection* in Wickens' model becomes *determine movement* in FIND. The decision where to move depends on the current *task.*

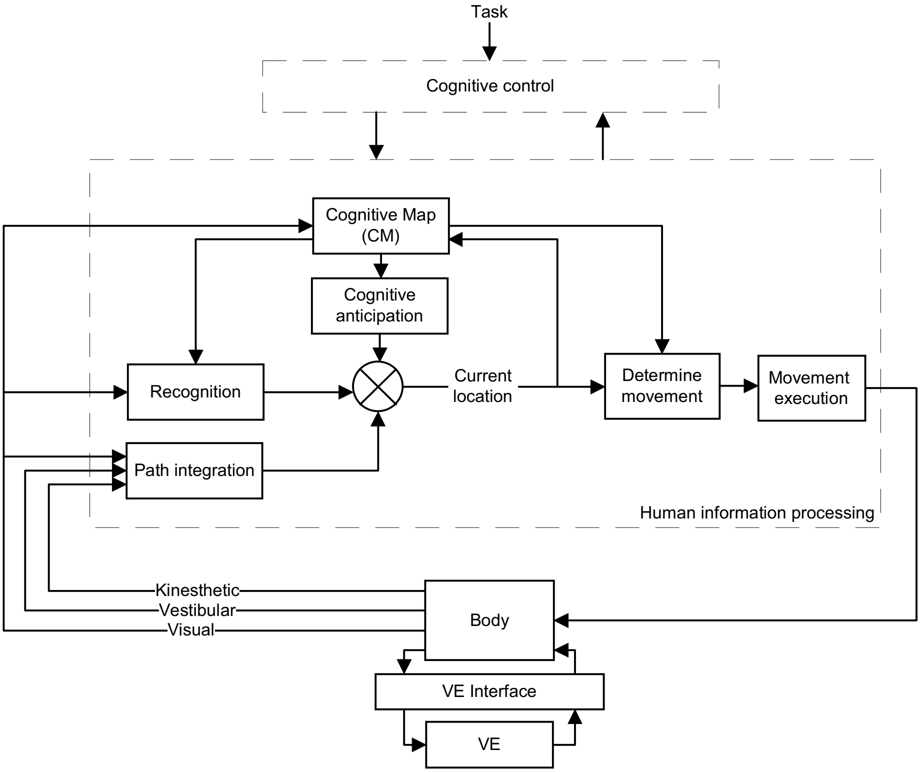


Figure 1.2: Framework for the Investigation of Navigation and Disorientation (FIND). The framework describes the mental processes and the information flows that are involved in spatial orientation in a VE. An operator interacts with a VE, mediated by the VE interface. In order to find information in the VE, an operator has to *determine a movement* and to *execute the movement*. To make this decision the operator needs to have knowledge about the environment that is stored in the *cognitive map*. Furthermore, the operator needs to know what is the *current location*. The current location is determined by a combination of three spatial updating processes: *path integration* that registers the displacement path through the environment on the basis of the available sensory feedback; *visual recognition* that uses the available knowledge in the cognitive map to recognise the location directly from the visual image; *cognitive anticipation* that uses knowledge about the environment and about the intended actions to determine the location, for instance after a discotinuous displacement. All processes are controlled by a single *cognitive control process* that distributes attention between the other processes, depending on what is the task.

The decision to follow a specific movement path depends on knowing where the information is, and it depends on knowledge about one's own current location, and on knowledge of a possible path between the two. Knowledge about the location of information and about possible paths is stored in memory. The spatial memory that is used to determine a movement is generally referred to as the *Cognitive Map* (CM). Kuipers (1982) warns about the use of the *map in the head* metaphor implied by the term *cognitive map*, because cognitive maps have many properties that do not correspond to geographical maps.

The CM can either be built through sensorial exploration of the environment, by learning from information sources like with geographical maps, or by inference from previous experience in similar environments (Thorndyke & Hayes-Roth, 1982; Darken & Sibert, 1996). The CM is defined in this thesis as the collection of all the information stored in memory that is useful for spatial orientation.

Apart from the CM, the operator needs to have some internal representation of his own *current location*, both the position and the orientation, in order to determine a route or a direction of movement. Let us now focus on how an operator determines his current location. Because the operator moves through the environment the internal representation of his current location must be repeatedly updated.

Three parallel processes can provide the information needed for this spatial updating:

1. *Recognition*: Recognition of the invariant structure of the environment depends primarily on visual perception, although hearing or scent may, in some cases, also provide direct information about one's location. In this thesis only visual recognition is considered. Recognition assumes familiarity with the environment. To be able to recognise your location in the environment a person needs to have a CM with which the perceived outside world can be compared. In the literature on spatial orientation, recognition of location is often associated with landmarks, which can be defined as any visual feature of an environment that can be associated with a specific place. Landmarks need to be unique to some degree to avoid confusion. The usefulness of a landmark for navigation purposes depends on whether the landmark can be seen from many places or from a long distance. Landmarks are used redundantly, that is to say, that a large part of the landmarks can be removed without performance being affected (Schenk, 1998; Steck & Mallot, 2000).
2. *Path integration*: Path integration means that the displacement during movement is registered by the integration of the available movement stimuli. If one knows the initial location and the displacement, one can determine the location after the movement. With natural locomotion the most important movement stimuli for registering displacement are the proprioceptive feedback picked up both by the kinesthetic senses and by the vestibular organ, and the optic flow sensed by one's eyes. Farrell and Robertson (1998) and Farrell and Thomson (1998) showed that path integration during locomotion without vision is automatic and requires no deliberate attention. This has not been proven for visual path integration.

   A well-known property of path integration systems is that they are subject to drift. Any systematic bias in motion measurement, however small, will eventually lead to a discrepancy between actual position and integrated position in the path integrator.

Although very accurate path integration systems have been demonstrated in animals who live in poor visual surroundings like for instance the desert ant or crabs (Healy, 1998), the human path integration system is much more inaccurate (Howard & Templeton, 1966, overview). The drift in path integration necessitates some form of reset of the process, from time to time, to compensate for increased error. Displacement during path integration is registered in relation to an initial location or some other reference location (Wan, Touretzky & Redish, 1993). Resetting could mean adjusting the integrated displacement to the observed displacement or could mean taking a new reference point and setting the displacement to zero. The reset is not explicitly modelled in FIND.

3. *Cognitive anticipation*: The combination of knowledge about one's original location and about the intended or real actions can be used to infer where we are or even where we will be in the future. Cognitive anticipation uses no feedback stimuli about the movement. Cognitive anticipation may possibly facilitate recognition, because it allows us to activate the appropriate part of memory beforehand. When placed in Wickens' model, *cognitive anticipation* would not be a part of *perception* but a part of *decision and response selection*.

Recognition, *path integration* and *cognitive anticipation* operate to determine the internal representation of our current location in the world. If this spatial updating fails, spatial disorientation occurs, which means that the internal representation of the location no longer corresponds to the actual location in the world. The severity of spatial disorientation depends on the effort or on the time needed to restore a correspondence between the internal representation and the actual location in the world. The *cognitive control* process divides the available resources between all the other processes. Cognitive control receives knowledge about the internal states of the processes and about the resulting behaviour. Going on this feedback, *cognitive control* may activate or suppress the execution of other processes.

1.2.2 The effects of interface technology on navigation behaviour

In a VE, the environment an operator perceives is created by the interface. The stimuli that are available for *recognition* or for *path integration* are not natural, but depend on the interface technology chosen. The execution of movement depends directly on the type of input-device that is provided. Other FIND processes have no direct relation to the interface, but might still be influenced indirectly. In the following paragraph, the differences between the two main types of VE-interfaces will be discussed. After that, the possible effects of these differences on the different FIND processes will be discussed. The effects of discontinuous navigation will be discussed separately in section 1.2.3.

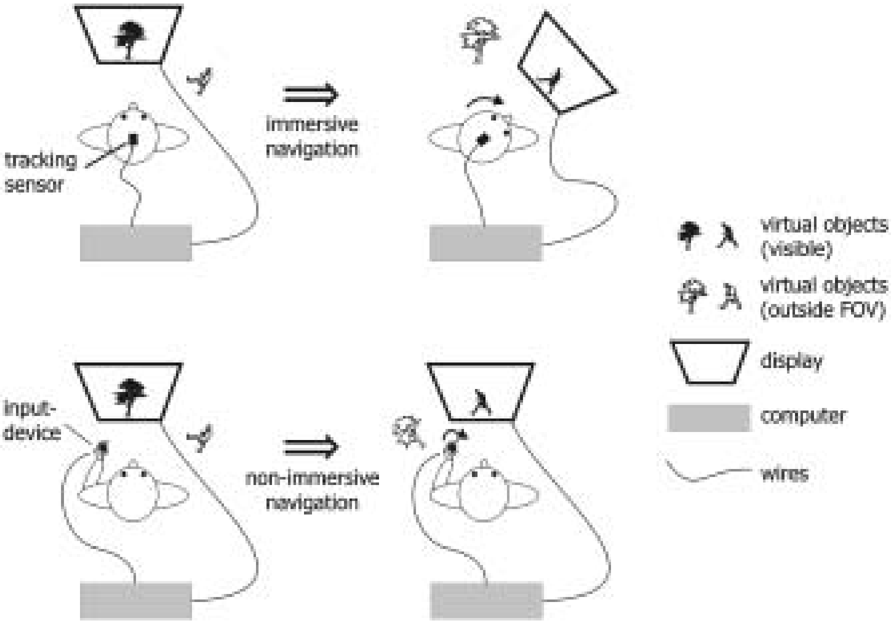


Figure 1.3: The principle difference between immersive and non-immersive navigation. On the left an observer is seen from above looking at a display on which a tree in the VE is shown. On the right the observer changes his viewpoint in the VE to face the walking man, making use of either an immersive interface or a non-immersive interface. With the immersive navigation interface (top), a natural movement is made by the observer who has a display that is fixed to the head. A head-tracking sensor registers the displacement, thereby providing input to the computer to render an image that corresponds to the changing viewing direction. With non-immersive navigation interface (bottom), the movement is controlled indirectly with a hand-held input-device while both the observer and the display remain stationary.

Immersive versus non-immersive navigation interfaces

Immersive navigation is only one aspect of the notion of *immersion* used to indicate the degree of sensory involvement in a VE, in a broader sense. The main difference between immersive and non-immersive navigation is illustrated in Figure 1.3:

- *Immersive navigation*. With immersive navigation, the movements of the virtual viewpoint in the VE are slaved to the head movements in the real world by using a head-tracking sensor. Natural head movements can be made in the real world causing

identical movements in the virtual viewpoint. Vestibular and kinesthetic feedback from the body can be used in a natural way to register displacements.
- *Non-immersive navigation*. With non-immersive navigation, the movements of the virtual viewpoint are controlled indirectly with some form of input-device. The operator remains stationary and only the image content moves, which means that natural feedback from the body is absent.

What is important to notice is that in both cases equal displacement through the VE will result in identical visual stimulus. The only difference is that with immersive navigation, extra information is available from both the vestibular organs and the kinesthetic senses from the different parts of the body. This feedback can be used for path integration to determine the displacement.

An immersive navigation interface is usually associated with an immersive display like an HMD, and a non immersive navigation interface is usually associated with a non-immersive display like a desktop monitor. However, this need not be the case. For instance, immersive navigation is used for the motion capture of actors in the animation industry, without the use of an immersive display. Also, many vehicle simulators provide an immersive display without immersive navigation. In this thesis, immersive navigation is investigated, whereas the type of display will not be varied.

Even with immersive navigation technology the movement in a VE is often restricted because of the limited range of most of the current tracking sensors or because of the limited available physical space for moving around freely. To alleviate this problem, advanced input-devices have been developed which enable movements that closely resemble natural locomotion to provide the input without leaving the spot. For instance, using an omni-directional treadmill (Iwata & Yoshida, 1999), or using a sliding surface on which users can walk without actual displacements (Iwata & Fujii, 1996), or simply stepping on the spot and using leg movement registration as input (Templeman et al., 1999). These devices offer some kinesthetic feedback, although not completely natural, and offer vestibular feedback for rotation but not for translations.

An advantage of immersive movement that has been demonstrated in a number of studies is that it improves depth perception (e.g. Smets, 1992; Jobling, Mansfield, Legge, & Menge, 1997; Voorhorst, 1998; or Barfield, Hendrix, & Bystrom, 1999). A disadvantage is that immersive VE has proven to be somewhat more nauseogenic than non-immersive systems, which is often attributed to the current limitations of head-trackers that create delays in the head-slaved image loop (Kolasinski, Goldberg, & Hiller, 1995; Howarth & Finch, 1999; Stanney & Kennedy, 1997; and McGee, 1998).

An important question to ask in the context of this thesis is whether immersive navigation, supported by these technological developments, ultimately contributes to improved performance in tasks requiring spatial orientation on the part of the operator. Looking at the FIND framework, we see that a direct influence of the vestibular and kinesthetic stimuli is only present for the path integration process. Therefore, immersive navigation can only improve performance in a navigation task if it improves the path integration process itself. Furthermore, if immersive navigation is to have an effect on navigation tasks, the path integration process needs to contribute to the internal representation of the *current location*, even in cases where visual recognition and cognitive anticipation are also possible.

In the following paragraphs, evidence from the literature will be considered to provide insight into the effects of the type of navigation interface on the different processes defined in FIND.

Movement execution

There can be no doubt, that the type of input-device that is provided to generate movement has a great effect on the possible speed and accuracy of the actual movement (e.g. Bowman et al., 1997; Breedveld, 1996; Lampton, Knerr, Goldberg, Bliss, Moshell, & Blau, 1994; Mackinlay, Card, & Robertson, 1990; Ware & Osborne, 1990; and Ware & Slipp, 1991). It is also clear that due to the limited speed of human locomotion, movement with non-immersive input-devices will always be able to outpace immersive movement. This might be different if the accuracy of movement response is important because then the additional proprioceptive feedback might become useful. Although the ergonomics of input-devices is an important issue, the aim of this thesis is to investigate the effect that the choice of interface technology may have on finding information. The ability to find information or to determine a movement depends on the ability to form a cognitive map and the ability to determine the current location.

Determine movement

Clearly, the difficulty of determining a movement to attain a current goal depends on both the quality of the information stored in the CM and on the accuracy of the internal representation of the current location. Because the latter also depends on path integration, the ability to determine a movement may depend indirectly on the type of navigation interface.

Interestingly, the access to information that is stored in the CM will depend on the current location. Sholl (1987) and Easton and Sholl (1995) showed that objects located in front of us are remembered more easily than objects located behind us. Therefore, determining a route from the current location to some other location is a totally different task from determining a route between two locations that do not correspond to the current location.

A pointing task can be seen as a basic form of determining a route between the current location and some other location. Knowledge is needed about the own current location in the environment and the location of the object if a response is to be given. An advantage of using a pointing task in an experimental setting is that it provides the possibility for determining response times which indicate the time needed for mental processing.

Darken and Sibert (1996) showed that when exploring a new VE many participants resort to systematic search patterns. Other heuristics are based on guidance provided by the available visual structure. For instance, when overlaying a visible grid over a VE, participants follow the gridlines during exploration. In the absence of a grid, participants frequently followed the coastlines of the islands that were simulated.

Cognitive Map

Properties of the CM have been studied since the first half of the twentieth century and it goes beyond the scope of this thesis to give an extensive overview of this (for recent overviews see Foreman & Gillett, 1997 and 1998). In the following paragraphs, some properties will be mentioned briefly.

The knowledge contained in the CM does not have to be integrated, but is built of disconnected components, with little or no relation between the components. There is substantial evidence that parts of the CM are hierarchically organised which leads to various distortions (Hirtle and Jonides, 1985; McNamara, 1986; and McNamara, Hardy, & Hirtle, 1989). Topological relations such as connectivity, order, and containment are represented, or at least retrieved and manipulated separately, from metrical relations of distance and direction (Kuipers, 1982; Montello 1991; McNamara, Ratcliff, & McKoon, 1984).

Spatial knowledge is best gained by active interaction using many different sensory modalities (Cohen, 1985). In the course of exploration the CM is built. Gaps in the CM can be filled by directing navigation towards unknown areas. Exploration guided by hiatuses in the knowledge contained in the CM might very well explain part of the advantage of active over passive exploration during spatial knowledge acquisition (Christou & Bülthoff, 1998; and Péruch, Vercher, & Gauthier, 1995). Passive observers do not get to see the specific locations they need in order to supplement their CM.

Péruch (1999) also found that a VE of a campus with distributed buildings is memorised more easily if the buildings are connected by visible roads that can be followed during exploration. This shows that the actual configuration and layout of a VE database can have a strong effect on spatial learning. However, this area will not be considered within the scope of this thesis, due to the limited amount of time available for this research.

Coding of the locations of objects in memory can be done with respect to different co-ordinate reference frames (Woodin & Allport, 1998). The most important distinction is to be made between egocentric and allocentric reference frames. Egocentric reference frames are *attached* to a part of the body, for instance to the retina, the head, or the trunk. Allocentric reference frames are attached to the environment.

Several authors have suggested that path integration might be especially important for the acquisition of a CM (Foreman & Gillett, 1998; McNaughton, Knierim, & Wilson, 1995; and O'Keefe & Nadel, 1978, p94). When exploring a new environment, recognition of the location is not yet possible because this requires a CM that is still missing. In order to encode the location of objects, path integration is needed to provide information about the relative location of objects. When two objects are close and can be seen in a single view the relative location of the two objects can be seen directly. If the two objects are further apart, one needs to move from one object to the other. Path integration provides an estimation of the magnitude of this movement and can therefore help to encode the location of two objects in relation to each other.

Research with immersive VE has shown that VE can be used effectively to build a CM of the spatial layout of a building, a ship, or a terrain (Bliss, Tidwell, & Guest, 1997; Darken & Sibert, 1996; Johnson & Stewart 1999; and Witmer et al., 1996) although VE training is still less effective than real world training (Bliss, et al., 1997; and Witmer et al., 1996). Similar results were found with non-immersive VE systems (Péruch et al., 1995; Ruddle, Payne, & Jones, 1997; Ruddle, Payne, & Jones, 1998; Waller et al., 1998; and Wilson et al., 1997).

Some recent studies make direct comparisons between immersive with non-immersive VE systems. Chance, Gaunet, Beall, and Loomis (1998) reported two experiments comparing immersive with non-immersive navigation. In the first experiment no significant difference in spatial updating performance was found between immersive navigation and non-immersive navigation, whereas in the second experiment immersive navigation was found to be advantageous. They give no explanation for this difference. Billinghurst, Bowskill, Dyer, and Morphett (1998) found a subjective preference among participants for immersive navigation to non-immersive navigation when looking around in a virtual information space, but they found no objective performance benefits. Ruddle, Payne, and Jones (1999) found some evidence that the metrical properties of a VE are remembered better with immersive navigation than with non-immersive navigation. Grant and Magee (1998) compare training of an exhibition floor between an immersive interface and a non-immersive interface. Although in the VE no difference in orientation performance was measured, the testing of knowledge transfer to the real world showed an advantage for immersive navigation. Waller et al. (1998) claim to compare immersive with non-immersive navigation but participants can control their viewpoint translations and rotations with a joystick in both the conditions. They found no difference in performance between the immersive and non-immersive conditions.

To summarise, the literature suggests that path integration plays a crucial role during the building of a CM. However, until recently no evidence was available to support this statement. Recent findings offer some support, but the evidence is still weak and some experiments do not show any path integration influence. The contribution of immersive navigation to spatial updating when visual recognition is also possible needs to be investigated in order to establish whether immersive navigation is beneficial for tasks requiring good spatial orientation.

Recognition

Recognition-based orientation may be harder in a VE than in the real world, because the relatively low number of objects and the low level of detail of the objects offered by most VE databases, contrasts sharply with the richness of visual detail available in the real world. One reason for having only a low level of detail is that there is a requirement for adequate rendering performance. Furthermore, modelling is time-consuming while copying virtual objects is fast and easy, often leading to the reuse of virtual objects, resulting in a VE where different locations look very much alike. Objects can then no longer be used as landmarks to determine the current location.

The degree to which path integration is involved in spatial updating when visual recognition or cognitive anticipation are also possible, is not clear. Logically, path integration is important in cases where cognitive anticipation and recognition are not possible or in cases where the movement is too fast to allow for recognition. However, with an abundance of visual cues in a familiar environment there might be no place for path integration because the current location can be determined by means of visual recognition. Attempts have rarely been made to show the role of path integration in the presence of abundant visual information.

Path integration

Path integration has been studied widely in both animals and humans (Maurer & Séguinot, 1995). Some animals have been shown to have very accurate path integration abilities (Healy, 1998). For instance, a desert ant searching for food runs straight home after finding a dead insect, even after wandering in a tortuous way for 600 meters. To show that the ant uses path integration and not visual cues, he is displaced laterally at the moment he finds the food. The straight path that the ant travels after this intrusion corresponds to the distance and direction to the home if he had not been displaced.

Path integration in humans can be performed even in situations in which not every sensory system provides information which would suggest that there is at least partial redundancy (Bles, 1981). Early research on spatial updating without visual recognition focussed on the contribution of the vestibular and the kinesthetic component. Evidence derived from investigating blindfolded participants showed that these cues could be used to determine the current location (Howard & Templeton, 1966, overview; or Loomis Klatzky, Golledge, Cicinelli, Pellegrino & Fry, 1993).

Sensory modalities have quite different characteristics, which may lead to different functional roles in path integration. The vestibular organ picks up acceleration information and becomes insensitive under sustained constant speeds (Mittelstaedt & Mittelstaedt, 1996; Bles, 1981). Therefore the vestibular system is particularly important for detecting motion onset, for discerning brief movement or for identifying frequently changing motions. The vestibular sense has distinct physiological organs for perceiving linear acceleration (otoliths) and angular acceleration (semi-circular channels) with quite distinct dynamic characteristics (Howard & Templeton, 1966; Bles, 1981; and Ivanenko, Grasso, Israël, & Berthoz, 1997).

The visual modality is slower at inducing a feeling of motion than the vestibular modality but it is able to detect speed. Visual flow patterns resulting from the different translations and rotations are quite distinct (Gibson, 1950; and Mulder, 1999) thus resulting in different sensitivities for different motions (Warren & Kurtz, 1992; and Klatzky, Loomis, Beall, Chance, & Golledge, 1998). With kinesthetic feedback, displacement of body parts can be registered directly without having to integrate velocity. A variety of receptors in our muscles (e.g. Muscle Spindles), tendons (e.g. Golgi Tendon Organ) and joints (e.g. Ruffini corpuscles) are responsible for registering forces and both the position as well as the velocity of our skeletal parts (e.g. Iggo, 1973).

There is considerable literature on the use of visual stimuli for self-motion perception. Warren, Morris, and Kalish (1988) define optical flow as "*…temporal change in the structure of the optic array, the pattern of light intensities in different directions at a moving point of observation.*" Gibson (1950) first described the optical flow patterns resulting from ego-motion. The flow pattern contains information about the movement itself and about the structure of the environment (Gibson, 1979). The optical flow field provides a larger degree of information about the environment than does a static optic array (Koenderink, 1986). Optical information supersedes kinesthetic or vestibular input when it comes to the specification of ego-motion (Gibson, 1979, Koenderink, 1986). Detection thresholds for heading directions are around one degree (Warren et al., 1988).

Path integration with only a visual stimulus in a VE has been investigated by Péruch, May and Wartenberg (1997) and by Van Veen and Riecke (1999). Péruch et al. (1997) found severe systematic overestimation of directional changes in a triangle completion task, depending on visual path integration. Van Veen and Riecke (1999) found only small errors in turning angles in a similar visual triangle completion task, but they provided their participants with extensive training before the experiment started. Witmer and Kline (1998) found evidence that for the perception of traversed distance a treadmill does not provide adantages over non-immersive navigation with a joystick, and that both suffer from a larger underestimation of traversed distance when compared to the real world.

To summarise, different sensory modalities can be used in path integration. The choice of VE technology determines which of these sensory modalities are stimulated during navigation. Evidence from the literature suggests that visual feedback alone might already be sufficient for path integration. However, some results indicate that path integration can be inaccurate if there is no proprioceptive feedback. There is insufficient evidence to support the notion that additional proprioceptive feedback, in addition to visual feedback, may further improve path integration performance.

Cognitive Anticipation

Cognitive anticipation does not depend directly on the VE-interface. However, this does not mean that no information from the environment is used. Signs indicating the destination of a discontinuous displacement need to be read. If the visual interface is so poor that reading becomes impossible then the cognitive anticipation that depends on these signs is no longer possible. However, generally this will not be the case.

Little is known about the dynamics of such cognitive anticipation in spatial updating. Neither has any information been found about the effectiveness of cognitive anticipation in spatial updating. Cognitive anticipation may be similar to mental displacement in which case a movement is not actually executed but merely imagined. Participants are asked to point at an object as if they were standing at a location that is different from their current location. This pointing generally takes longer and is more error prone than pointing from the actual location. The mental displacements need more time as the magnitude of displacement increases (Boer, 1991; Easton & Sholl, 1995; Gaunet, Martinez, & Thinus-Blanc, 1997; Presson & Montello, 1994; and Rieser, 1989). Whether these findings also apply to displacements that are actually made after cognitive anticipation is not clear. If cognitive anticipation was similar to mental displacement, then the results found with mental displacement would predict that cognitive anticipation may require high mental effort and may lead to an increase in error, when compared to other modes of spatial updating.

Cognitive control

Cognitive control may play an active part in spatial updating. There are three parallel processes of which the information need to be combined into a single internal representation of the current location. The integration of the information from *path integration*, *recognition*, and *cognitive anticipation* cannot be seen as a simple summation. There may be many interactions between the processes in FIND that are not explicitly depicted with arrows in the diagram (Fig.1.2). For instance, the cognitive anticipation of a discontinuous displacement may facilitate recognition after the displacement. The outcome of the visual recognition process may possibly be used to reset the path integration process.

How the different sources of information are combined will depend on the task and on an evaluation of the validity of the information provided by the processes. Depending on the task setting, cognitive control may suppress the execution of one or two of the three processes. Whether this is possible without incurring a decrement in performance may also depend on the degree to which the process in question is automated.

1.2.3 The effect of discontinuous displacement

So far, the role of immersion in spatial orientation in VE has only been considered in relation to continuous navigation. As discussed in the background section, there is also a need for more efficient discontinuous displacement. The clear advantage of discontinuous displacement is that it can be extremely fast. However, the jump in location may cause a discrepancy between a human's internal representation of the current location and the actual location in the environment. This discrepancy was defined as disorientation. The severity of this disorientation is measured by the time or effort needed to restore a correspondence between the internal representation and the actual location. Firstly, discontinuity will be defined and after that the possible effects of discontinuous navigation will be discussed.

The distinction between continuous and discontinuous displacement is partly artificial. Because of the limited update-rate of the images in a VE system, every displacement is by definition discontinuous. Things may be termed continuous when the discontinuities become extremely small. On the one hand, if the update-rate of the images is sufficiently large and if the magnitude of the displacement is sufficiently small, the sequence of transitions between subsequent images will be perceived as fluent by the human observer. On the other hand, as the frame-rate drops or the displacement between subsequent updates of the image increases, the coherent senses of motion will break down (Spillmann & Werner, 1990). The precise value of the breakdown boundary depends on several factors, like for instance the spatial frequency of the stimulus, and the number of consecutive frames that are shown. Breakdown of motion does not mean that a recognised object can no longer be associated with the same object when it is recognised in a different location one frame later. With similar objects, aliasing may occur, leading to apparent motion like the camera images that make a vehicle tire appear to revolve in the opposite direction.

Although the breakdown of motion sensation could be taken as a possible boundary between continuity and discontinuity another approach can be taken based on the overlap of image contents between two successive frames. During the display of one frame, the area of the display surface is occupied by a limited set of object surfaces. When the observer moves, object surfaces move in and out of view and grow larger or smaller. The percentage of the display surface that is occupied by object surfaces that were also present in the previous frame is defined as *visual overlap* between the two frames. Continuity may then be defined by the amount of visual overlap.

Total discontinuity means that none of the object surfaces that could be seen in the original frame can be seen in the subsequent frame. According to this definition there are two ways in which total discontinuity can occur. Object-surfaces can move out of the Field Of View (FOV) or they can be occluded by other surfaces. For rotations of the viewpoint larger than the FOV, there are no surfaces that can be seen in both the subsequent images. For translations of the viewpoint, at least some occlusion is needed to gain total discontinuity.

Discontinuous displacement may lead to disorientation, which means that there will be a discrepancy between the internal representation of the current location and the actual location. Because there are no movement stimuli, the registered displacement by *path integration* during the movement is zero. The internal representation of displacements as registered by path integration, does not therefore, correspond to the actual displacement. Besides that, the internal representation of the current location is still based on the previously recognised view. The new view, after discontinuous displacement, needs to be recognised in order to update the internal representation of the current location. The recognition can be seen as a search process in the CM for a corresponding location. This search may become less difficult depending on whether or not the new location can be anticipated. If the destination of a discontinuous displacement is known, then cognitive anticipation may be able to compensate beforehand for the loss of movement feedback.

The disruptive effect of a discontinuous displacement may well depend on some characteristics of the displacement. For instance, the spatial relations between the start-point and the end-point of the displacement may have an effect on how much interference is caused by incorrect information from path integration. The functional relatedness between the start-point and end-point may have an effect on cognitive anticipation as well.

There is very little literature on the effects of discontinuous displacement in three-dimensional environments. Bowman et al. (1997) found evidence to support the notion that discontinuous displacement leads to temporary disorientation. They investigated the spatial updating of participants who were translated (passively) either discontinuously or continuously to a location in the viewing direction of a known VE. Their results showed that participants had to search significantly longer for targets (approximately one second) after a discontinuous jump than after a continuous movement to their final position. However, it is unclear whether participants were provided with sufficient information to use cognitive anticipation effectively.

If discontinuous displacement leads to temporary disorientation then spatial knowledge acquisition may also be disrupted. Experiments by Witmer et al. (1996) showed that it is possible to learn a building layout from route directions in combination with pictures of landmarks. The act of viewing a series of pictures can be seen as discontinuous navigation. Therefore, this result shows that participants are able to integrate the information that is gained from discontinuous views into a coherent CM, at least with the additional verbal directions that were given. However, learning a building's layout by using a VE with continuous navigation proved more effective.

Péruch et al. (1995) found that participants had a subjective preference for continuous displacement as opposed to discontinuous displacement during exploration of a VE. They investigated spatial knowledge acquisition in a maze, by comparing animated continuous movement at 18 frames per second with discontinuous movement at one frame every four seconds along similar pathways. The movement for both the conditions was passive which meant that the participants did not control their movements themselves. Although no differences in performance measures were found between the two conditions, a majority (70%) of the participants indicated that discontinuous movement demanded more attention. Similarly, Billinghurst et al. (1998) also found a subjective preference for continuous displacement but again no objective difference in performance was found.

To summarise, little is still known about the effects of discontinuous displacement on spatial orientation behaviour. Some evidence was found that discontinuous displacement leads to temporary disorientation. It is unclear whether cognitive anticipation may have been able to compensate for this disorientation. Some subjective evaluations indicate that the acquisition of spatial knowledge may be harder when navigating with discontinuous displacements but no objective evidence was found to support this. None of the studies found systematically investigate the role of cognitive anticipation, nor was any research found on different types of discontinuous displacement.

1.2.4 The limitations of current technology

Several limitations of current VE technology may have an effect on spatial orientation. Although these limitations are not the focus of study in this thesis, the most important technical parameters that do ave an effect on performance will be briefly summarised.

Latency

With current VE technology, the latency between the measurement of the head position and the subsequent updating of the graphical image may have detrimental effects on performance (Frank, Casali, & Wierwille 1988). The rendering delay is usually proportional to the number of polygons (Akatsuka & Bekey, 1998), which means that a trade-off exists between visual detail and rendering performance.

Latency changes the eye movements. When turning under natural circumstances, eye movements are generated reflexively in response to vestibular feedback in order to stabilise the retinal image and avoid blurring. This is what is known as the *vestibular-ocular-reflex.* Natural turning usually starts with a fast saccadic eye movement in the direction of the movement followed by slower head movement. While the head is moving, the eyes re-centre in their sockets and stabilise in space with the aid of the vestibular-ocular-reflex (Land, 1992).

With an immersive VE interface visual feedback is delayed. When starting to move the head and the HMD with it, the image is not yet updated and it remains unchanged. This results in an instability of the virtual image which means that the virtual objects displayed shift in relation to the intended location. At the same time, the eyes make a movement triggered by the vestibular-ocular-reflex to compensate for the head movement. The result is that the eyes move relative to the objects in the image instead of keeping focussed on it, which results in image blur (Barnes & Sommerville, 1978; and Sandor & Leger, 1991).

A possible strategy for coping with the effect of the delays is to move more slowly, thereby minimising the experienced discrepancy (De Vries & Padmos, 1997; Wells & Venturino, 1990). Several investigations have shown that the delays in a head-slaved image generation loop reduce performance, for instance in tracking tasks (So & Griffin 1993), driving tasks (Padmos, 1999), or when searching for targets (Van Erp & Van den Dobbelsteen, 1998a and 1998b).

In a non-immersive VE the pattern of eye movements adopted to stabilise the retinal image is different. Manipulating the input-device generates turning, and the eyes have to pursue the visual stimulus in order to stabilise the retinal image without using vestibular signals. Fast saccadic eye movements are needed from time to time to re-centre the eyes. The difference when compared to natural turning is that the slow and fast movement phases are exchanged. Furthermore, the vestibular signal is absent, which means that there is no involuntary eye movement due to the vestibular-ocular-reflex. The slow movement is then generated by the eye pursuit mechanism, which has different dynamic characteristics than the vestibular-ocular-reflex (Barnes & Sommerville, 1978; Griffin, 1990; Moseley & Griffin, 1986; Paige & Seidman, 1999; and Sandor & Leger, 1991).

Field Of View

In most displays, the viewing angle in the VE or Field Of View (FOV) is restricted. This limits the number of objects that can be seen simultaneously, thereby also limiting the chances of correct recognition. A limited FOV has a detrimental effect on various tasks like for instance target detection (Osgood & Wells, 1991), tracking (Kenyon & Kneller, 1992), or helicopter manoeuvring (Edwards, Buckle, Doherty, Lee, Pratty, & White, 1997). Alfano and Michel (1990) showed that with a highly reduced FOV of nine degrees participants could no longer correctly judge where they were in relation to the room and so they experienced bodily discomforts like unsteadiness and dizziness. McConkie and Rudmann (1998) found a reduced accuracy of the CM for a FOV that was restricted to 19 degrees when compared to larger FOVs of 30 degrees and higher. Even with less restrictive FOVs, McCreary and Williges (1998) found performance decrements in cognitive mapping tasks comparing a 30 degrees with a 48 degrees FOV. Wells, Venturino, and Osgood (1988) and Venturino and Kunze (1989) found no effects from reducing the FOV down to 20 degrees on the accuracy of the CM. Johnson and Stewart (1999) also found that a reduced FOV of 40 degrees has no effect on the acquired spatial knowledge.

Resolution

Poor resolution hinders recognition. If resolution is to be increased, more expensive displays will be required and an increase in rendering power will be demanded. With some special applications very low resolutions can still be found, for instance in teleoperation systems with limited bandwidth communication (Van Erp & Van den Dobbelsteen, 1998a). Smets and Overbeeke (1995) have shown that poor resolution can be partly compensated for by enabling head-slaved movement to take place. Increasing the resolution also increases the demands on the rendering engine. A smart, though technically more demanding solution, is to place a high resolution insert in a low resolution surrounding thus combining the demand for a large FOV with the demand for high foveal resolution (Kappé, 1997; and Yoshida, Rolland, & Reif, 1995). When using an HMD, this solution requires additional eye-tracking to generate input for the displacement of the high resolution insert, whereas when using a large screen display a head-tracker may be sufficient.

Accommodation-vergence mismatch

With current commercially available stereoscopic HMDs, the image focus plane of the optic set-up is fixed. This leads to a fixed required distance for eye accommodation, whereas eye vergence is determined by the location of the object in depth as required by the image disparities (Rushton & Riddell 1999). This conflict between accommodation distance and vergence distance is not present under normal viewing conditions. The cue conflict leads to incorrect depth perception (Edgar, Pope, & Craig, 1994). Furthermore, serious problems were found with HMDs leading to eyestrain (Edgar et al., 1994; Hasebe, Oyamada, Ukai, Toda, & Bando, 1996; Kawara, Ohmi, & Yoshizawa, 1996; and Mon-Williams, Wann & Rushton, 1993). These problems are not present when accommodation and vergence are adequately matched (Mon-Williams & Wann, 1998; and Rushton, Mon-Williams, & Wann, 1994). New developments in HMD design promise to solve these problems in the future, but no commercial versions of these HMDs are yet available (Onishi, Yoshimatsu, Kawamura, & Ashizaki, 1994; Shiwa & Miyasato, 1997; and Sugihara & Miyasato, 1998).

To summarise, the visual stimulus in a VE is less than natural because of current technical limitations, which may hinder recognition. Path integration might be hindered by limited update rates and latencies.

1.3 Questions and Hypothesis

As explained in the background section, there is a need to understand the potential advantages and drawbacks of using immersive versus non-immersive VE technology as an interface to support human operators.

The framework FIND shows that immersive-navigation changes the information available for path integration. If the choice of VE technology influences spatial orientation performance, this should be noticeable from the difference in path integration performance. Evidence from the literature suggests that visual feedback alone might be sufficient for path integration. However, some results indicate that path integration can be inaccurate without proprioceptive feedback.
The first research question and hypothesis are:

Q1. When compared to non-immersive navigation, does immersive navigation improve the quality of path integration?

H1. An interface that allows immersive navigation can improve the quality of path integration when compared to an interface that only allows non-immersive navigation.

Even if H1 is true, path integration is still only one of three possible alternative processes that can be used in spatial updating. Any possible effect on the quality of path integration might be overshadowed by the contribution that visual recognition or cognitive anticipation makes to the determination of the current location. Suggestions have been found in the literature to the effect that path integration should play a part during the acquisition of a CM. However, evidence to prove this suggestion is still very weak. Hence the second question and related hypothesis which is:

Q2. When compared to non-immersive navigation, does immersive navigation improve the acquisition of a cognitive map?

H2. Immersive navigation can improve the acquisition of a cognitive map.

The first two questions focus on the effects of the choice of a type of interface. As well as having continuous displacements, discontinuous displacements will also be necessary if efficient navigation over larger distances is to be achieved. Little is known about the effects of these discontinuous displacements and how path integration, visual recognition and cognitive anticipation interact. So, the logical ensuing question is:

Q3. Does discontinuous displacement affect spatial updating?

H3. Discontinuous displacement can disrupt spatial updating.

This third question is more broad than the first two questions and it will be split into three sub-questions. If immersive navigation has an effect on path integration (H1) and on the acquisition of a cognitive map (H2), this shows that spatial updating partly relies on path integration. However, with discontinuous displacement, path integration provides incorrect information or even no information due to the lack of movement feedback. Little is known about the effects of discontinuous displacement on spatial knowledge acquisition. Hence, Q3 and H3 will be redefined in the form of three detailed questions and hypotheses:

Q3a. Does discontinuous navigation impair the acquisition of a cognitive map?

H3a. Discontinuous navigation can impair the acquisition of a cognitive map.

If H3a is true, what is the cause of performance decrement remains unclear. If the right information is provided, cognitive anticipation could in principle compensate for the lack of correct path integration information. The efficiency of the processes may depend on how much they are automated by repeated experience.

Q3b. Is cognitive anticipation able to compensate for the disruption of spatial updating that is caused by discontinuous displacement?

H3b. Cognitive anticipation can reduce disorientation caused by discontinuous displacement.

If discontinuous displacement disrupts spatial updating, the severity of this disruption may well depend on the type of displacement, which is determined by certain characteristics like for instance the spatial relationship between the start-point and the end-point of the displacement.

Q3c. Does the type of discontinuous displacement have any effect on spatial updating?

H3c. The type of discontinuous displacement can have an effect on spatial updating.

Finding a type of displacement effect would mean that guidelines could be given for the type of discontinuous displacements in a VE interface.

1.4 Outline of the thesis

To answer the questions formulated here, a series of experiments was executed. The experiments have been grouped, according to the similarities in their task and stimulus material, into four chapters (Chapters 2 to 5).

In Chapter 2, two experiments are reported which focuss on Question 1. In Experiment 1, the quality of path integration is investigated with different combinations of visual, vestibular, and kinesthetic feedback. In Experiment 2 the use of visual feedback for path integration is further investigated.

In Chapter 3, a series of three experiments is reported further elaborating on Question 1, which investigated whether path integration with a non-immersive interface can be trained so that the same performance as with an immersive interface can be obtained. The first of these, Experiment 3, investigates whether explicit feedback on the performance can help to train path integration that is based only on visual feedback. Experiment 4 is a control experiment that checks whether path integration performance is influenced by characteristics of the visual stimulus other than optic flow. In Experiment 5 training of path integration is tested with implicit feedback that is provided by enabling recognition of the surrounding scene to take place.

The remaining four experiments reported in Chapter 4 and Chapter 5 each consist of at least two parts. In the first parts that focus on Question 2, spatial knowledge acquisition is tested with different navigation interfaces. In the second parts that focus on Question 3, spatial updating of the previously learned layout is tested after different types of continuous and discontinuous displacement have occured.

In Chapter 4, two experiments are reported in which a limited set of objects had to be learned. In Experiment 6, the testing of learning with an immersive and a non-immersive navigation interface is described both with continuous and with discontinuous navigation. The second part of the experiment verifies whether discontinuous displacement leads to disorientation when cognitive anticipation is not possible. In Experiment 7, learning with an immersive and a non-immersive navigation interface is tested again. Besides that, the effect of training path integration beforehand on learning with a non-immersive interface is investigated. In the second part of this experiment the spatial updating is tested not only during known displacement and after random displacement, but also after no displacement has taken place, thus making it possible to compare cognitive anticipation on the one hand with recognition and path integration on the other hand.

In Chapter 5, the layout that has to be learned by participants is expanded. This allows for more different types of discontinuous displacement, thereby making it possible to investigate the dynamics of the spatial updating processes during discontinuous displacements. The two experiments reported are Experiment 8 and Experiment 9 which both use the same VE. The first parts of the experiments are identical and that is where the duration of spatial knowledge acquisition is tested with an immersive navigation interface or a non-immersive navigation interface. In the second part of Experiment 8 the dynamics of spatial updating is investigated during discontinuous displacement and in the absence of cognitive anticipation.

In Experiment 9, after learning the layout, spatial updating is compared for both continuous and discontinuous navigation with and without cognitive anticipation of the destination.

Chapter 6 provides the conclusion and a discussion in which the findings are related to the formulated research questions. The chapter ends with recommendations for further research and a summary of the implications of the findings.

2

Path integration[1]

Abstract

In two experiments, Hypothesis 1 is tested: An interface allowing immersive navigation increases the quality of spatial updating by path integration.

In Experiment 1, participants were put in a virtual forest and were asked to turn specific angles with navigation interfaces that provided different combinations of visual, vestibular, and kinesthetic feedback (pure visual, visual plus vestibular, visual plus vestibular plus kinesthetic, pure vestibular, and vestibular plus kinesthetic). Furthermore, in Experiment 2 the visual flow was manipulated by providing a 60% zoom for the same task.

The results confirmed Hypothesis 1, showing that kinesthetic feedback provides the most reliable and accurate source of information for path integration. Orientation on the basis of visual flow alone is most inaccurate and unreliable. In all the conditions, participants overestimated their turning speed and consequently did not turn far enough. Both the absolute errors in path integration and the variation in path integration increase as the path length increases.

2.1 Introduction

If a movement is continuous, the path covered can be integrated from the motion sensations provided by the *vestibular*, *kinesthetic*, and *visual* feedback stimuli of the body. Whether vestibular and kinesthetic information is available depends on the VE technology that is used. Although, the parallel sources of information suggest some redundancy, the vestibular and kinesthetic information might not be superfluous for path integration.

The goal of the two experiments reported in this chapter was to investigate the relative contributions that vestibular, kinesthetic, and visual feedback make to path integration in

---

[1] Parts of Chapter 2 have been published as:

Bakker, N.H., Werkhoven, P.J., & Passenier, P.O. (1998). Aiding orientation performance in virtual environments with proprioceptive feedback. *Proceedings of the IEEE Virtual Reality Annual International Symposium (VRAIS'98), Atlanta, GA*, pp.28-33.

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

VE. In the first experiment, the relative influence of the different feedback modalities is determined by using five navigation interfaces with different combinations of visual, vestibular, and kinesthetic feedback. A second experiment was carried out to further investigate the role of visual flow and to verify whether participants used alternate visual strategies like counting trees to determine the angle turned.

If we are to investigate path integration the contribution of visual recognition to spatial updating needs to be excluded. As yet, it is not investigated how path integration affects learning of the environment.

Navigation interfaces in VE

In the ideal case, a VE would provide sensations that are indiscriminable from real world experiences. In current technology we are still a long way away from this ideal. The user's movements in VE are usually restricted, both by the limited range of most current tracking sensors and by the limited available space for moving around freely. Some kind of input-device is therefore needed to enable movement over large distances. A common characteristic of these devices is that the user does not translate through the real world. The user therefore lacks certain feedback information that would normally be present during motion in the real world.

A large variety of hand-controlled input-devices are commercially available, allowing control of all six degrees-of-freedom such as, for instance, the SpaceMouse™, the DataGlove™, and the SpaceBall™ (for evaluations see Bowman et al., 1997; Ware & Osborne, 1990, or Zhai & Milgram, 1993). However, these input devices do not provide the vestibular and kinesthetic feedback that is available with natural locomotion.

Advanced input-devices have been developed that enable movement close to natural locomotion as input. Movements of the legs is made possible by using a treadmill (Iwata & Yoshida, 1999; and Witmer & Kline, 1998), a fitness bicycle (Riecke, 1998), a sliding surface (Iwata & Fujii, 1996), or simply by stepping on the spot (Templeman et al., 1999). By tracking the displacement of the movement device or of the legs, translation through the virtual environment is generated without actual displacements in the real world. Iwata and Fujii (1996), for instance, developed an input-device, in which the user performs walking movements on a low friction surface while being restricted to one position by a hoop at waist height. By tracking the position of both feet, input is generated to control movement through the VE. Templeman et al. (1999) measured knee movements and ground reaction forces while stepping on the spot to generate the virtual translations. Witmer and Kline (1998) used a treadmill for virtual translation in the forward direction. Iwata and Yoshida (1999) developed a bi-directional treadmill that allows virtual locomotion in a horizontal plane. Another approach that provides less feedback is that of letting the user lean over into a direction, which then generates a translation velocity in that direction. To generate the input, either the displacement from a centre (Wells, Peterson, & Aten, 1997) or the ground reaction forces are measured (Peterson, Wells, Furness, & Hunt, 1998).

The input-devices mentioned differ from the point of view of the kind of feedback that is available during movement. All devices lack vestibular feedback for translation. Some provide adequate vestibular feedback for rotation and all provide some form of kinesthetic feedback, although this is not completely natural. To choose between alternative navigation devices, knowledge is needed about the contribution of the different feedback modalities to path integration.

Sensory modalities for path integration

Although there may be some redundancy, the sensory modalities have quite different characteristics. This may lead to differences in the contribution to path integration for the sensory modalities. Vestibular organs pick up acceleration information and become insensitive under sustained constant speeds (Bles, 1981; and Mittelstaedt & Mittelstaedt, 1996). The vestibular system is therefore particularly important for detecting an onset of motion, for discerning short motions or for sensing frequently changing motions. The vestibular sense has distinct physiological organs for perceiving translations (otoliths) and rotation (semi-circular channels) with quite distinct dynamic characteristics (Bles, 1981; Howard & Templeton, 1966; and Ivanenko et al., 1998). The visual modality is slower than the vestibular modality to generate a sensation of movement, but it is able to detect speed. Visual flow patterns resulting from the different translations and rotations are quite distinct which leads to different sensitivities for different motions (Warren & Kurtz, 1992). With kinesthetic feedback, the displacement of body parts can be registered directly without having to integrate velocity.

Kinesthetic + vestibular feedback

Some animals that are unable to rely on visual recognition have very accurate path integration abilities, like the desert ant or certain species of crabs (Healy, 1998). Human path integration is less accurate (Howard & Templeton, 1966), possibly as a result of our reliance on visual recognition. The role of the kinesthetic and vestibular modalities in path integration, have long been compared by letting participants walk blindfolded, or by moving them around in some form of vehicle, for instance a wheelchair.

Sholl (1989) showed that path integration based only on vestibular information degraded as the number of path-segments traversed increased. After being wheeled around blindfolded in a wheelchair over a path with three segments, participants could no longer point to their point of origin. When participants were allowed to walk, they could point to the origin even after traversing the longest four-segment path. Israël, Sievering and Koenig (1995) investigated path integration only using vestibular feedback for rotation around the vertical axis. Participants showed undershoot errors with an average of 5.5% for instructed angles of 90Ñ, 180Ñ and 360Ñ. In a similar experiment, Israël, Bronstein, Kanayama, Faldon, and Gresty (1996) showed an earth-fixed target before the light was turned off and rotation began. This visual target reduced variability in responses showing that a visual reference position might also be useful for vestibular path integration. With walking conditions, several authors showed accurate performance. For translations up to 14 meters, Loomis, Da Silva, Fujita, and Fukusima (1992) showed that participants can walk accurately without vision to previously viewed targets with errors below 10 percent. Klatzky, Loomis, Golledge, Cicinelli, Doherty, and Pellegrino (1990) investigated how accurately participants could produce a turn that was displayed with the aid of a wooden clock, with natural walking but no vision. The errors do not increase monotonically for instructed angles although that would be expected if path integration was biased. Mean error ranges from 5Ñ to 30Ñ. Smaller errors were found for the orthogonal angles which would suggest that some cognitive mechanism is involved in their task. Sadalla and Montello (1989) found similar magnitudes of errors when participants are required to walk blindfolded along a pathway with one turn. The experiments reported so far do not quantify the role of kinesthetic feedback on its own without vestibular feedback.

To investigate kinesthetic feedback without vestibular feedback the supporting ground plane has to move. Witmer and Sadowski (1998) compared natural walking in the real world with walking on a treadmill in a VE both without vision, to a previously viewed target. Although they found larger errors in VE, it is not clear whether it is caused by the differences in visual perception of the target distances between the real world and the VE, or whether this is caused by differences in the perception of traversed distance between natural walking and walking on the treadmill.

Visual feedback

Extensive research has been done into heading direction perception during translation, with visual stimuli consisting of random dot patterns (Beintema, 2000, overview; Royden, Banks, & Crowell, 1992; Warren et al., 1988; Warren, Blackwell, Kurtz, Hatsopoulos, & Kalish, 1991; and Warren & Kurtz, 1992). Heading can be accurately perceived from the flow pattern generated by the moving dots, but in some cases the information in the flow pattern is ambiguous and heading is not perceived correctly. In such cases extra-retinal information is needed (e.g. efference copy, vestibular or kinesthetic neck muscle signals) to disambiguate the percept (Beintema, 2000; and Royden et al., 1992). Although correct heading perception might be a prerequisite for visual path integration these results do not show that the instantaneous heading information is indeed integrated to obtain a movement path.

Bertin, Israël, and Lappe (in press) investigated whether participants could reconstruct their path after passive optic flow stimulation. Results showed that the path shape is perceived correctly as long as viewpoint orientation remains tangential to the path but is otherwise perceived incorrectly. Péruch et al. (1997), investigated path integration only relying on visual feedback using a triangle completion task. They found severe systematic overestimation of directional changes. However, because in some of their experiment conditions a visual overview of the triangle is present, participants can recognise the direction and length of the final leg of the triangle beforehand. Therefore, this experiment does not distinguish clearly between spatial updating based on visual recognition and on path integration. In contrast to Péruch et al. (1997), Van Veen and Riecke (1999) found only small turning angle errors in a similar visual triangle completion task in VE, but they provided their participants with extensive training before the experiment started.

Warren and Kurtz (1992) reviewed the literature on the perception of motion from visual flow either using central or peripheral vision. Unlike previous notions to the effect that perception of self-motion is primarily based on peripheral vision, they conclude that central vision can accurately extract information on self-motion from visual flow. This means that the perception of velocity from visual flow should not be affected by the limited FOV offered by an HMD. It is interesting in this respect to observe, that if a zoom function is provided (Péruch et al., 1997), the visual overview and thus visual recognition can be improved. However, this zoom also alters the amount of visual flow that is available while moving around, which might have a negative effect on path integration.

Depth information provided by stereoscopic images is also important for visual path integration. Wist, Diener, Dichgans, and Brandt (1975) found that perceived rotary self-motion depends on the perceived distance of the visual stimulus. Palmisano (1996) found that stereoscopic images facilitate the onset of vection from optic flow.

Vestibular + kinesthetic + visual

Witmer and Kline (1998) found no main effect of interface type, when they compared a treadmill interface to a joystick interface in order to estimate the traversed distance in the forward direction. However, visual recognition of the expansion of the end of the virtual hallway in which the participants moved might have reduced the influence of improved path integration with the treadmill. Iwata and Yoshida (1999) also found no difference in the reproduction of translation distance in a VE between a bi-directional treadmill and a joystick interface. However, if the path included a curve, the bi-directional treadmill proved to be superior to the joystick.

Klatzky et al. (1998) found that participants in a VE did not update their mental representation of their orientation at all in a triangle completion task if shown an animation, whereas physical turning ensured proper spatial updating.

In experiments done in the real world with a rotating drum, Bles (1981) found that stimulating any of the seven possible combinations of the three sensory modalities can produce equal turning sensations for rotations about a vertical axis. With the purely vestibular condition the turning sensation ceases after a few turns.

To summarise, the results found in the literature indicate that active walking which provides both vestibular and kinesthetic feedback, leads to more accurate path integration than being wheeled around passively with only vestibular feedback. This is not surprising if one thinks that the vestibular apparatus can only detect acceleration but that it fails to detect constant velocity. Results regarding path integration in the presence of a visual stimulus are less clear. On the one hand, some results in VE suggest that the visual feedback cannot be used with great accuracy for path integration (Péruch et al., 1997; Klatzky, et al., 1998; and Iwata & Yoshida, 1999). On the other hand, however, some results suggest that the visual stimulus can be adequately used for path integration in the real world (Bles, 1981), or in a VE (Van Veen & Riecke, 1999). However, the usability of the visual stimulus depends on the specific motion pattern involved (Beintema, 2000; Bertin et al., in press). For pure translation in the viewing direction, the visual stimulus seems sufficient and additional feedback does not improve performance although the evidence is weak (Iwata & Yoshida, 1999; and Witmer & Kline, 1998).

Direct comparisons between path integration performance under the influence of different combinations of sensory modalities are scarce and seem contradictory. On the one hand, some results suggest that there is complete redundancy between the modalities, at least *for pure rotation (Bles, 1981). On the other hand, some results show that path integration performance is better if alongside of visual feedback, kinesthetic feedback is also present (Iwata & Yoshida, 1999).*

To investigate the role of vision in pure path integration an experimental task is needed, in which no overview is present and recognition plays no part. Although the triangle completion task has been widely used to investigate path integration (Klatzky et al., 1990; Lederman, Klatzky, Collins, & Wardell, 1987; Loomis et al., 1993; Péruch et al., 1997; and Sadalla & Montello, 1989) a number of fundamental problems limits its diagnostic value. Since translation and rotation are not independent, an error in perceived translation, despite a correctly perceived first rotation, will still result in an error in the executed second angle. Therefore, no clear distinction can be made between errors in translation and in rotation. In general, this and other path completion tasks (including pointing to an unseen starting point) may be regarded as consisting of at least two phases. In the first phase, the path is integrated from perception while movement is guided, for instance, by a visual target or by the experimenter. In the second phase, the movement to the starting point is executed on the basis of the integrated path and according to the available feedback information. The measured error is a result of flaws in perception in both the movement phases. Depending on the geometry of the path used, errors in the first phase may be cancelled by errors made in the second phase, thus leading to a small measured error, whereas large perceptual errors were made.

To avoid the problems mentioned, a task was chosen for the experiments, in which the CM is already well-known and accurate. It is well accepted that people have an accurate internal representation of an orthogonal reference frame, corresponding to the front-back and left-right symmetry of the human body (Yungkurth Hooper & Coury 1994). The orthogonal reference frame can be regarded as a CM that reliably indicates the four cardinal directions in relation to the body. Participants were asked to turn to this imaginary target and the error in the angle produced was measured. Since the CM used is accurate, the errors that are measured will only be caused by inaccuracies in path integration during movement. Because only one movement phase is present in the current task, the above-mentioned problem of cancellation of errors is avoided. Only rotation is investigated to avoid interaction with translation. Furthermore, translation is not investigated because the means for tracking translations outside a very limited volume were not available.

2.2 Experiment 1: Determining the influence of visual, vestibular, and kinesthetic feedback

What is measured in Experiment 1 is the accuracy with which participants can turn over prescribed angles corresponding to the orthogonal reference frame, in the presence or absence of visual feedback, using three different navigation interfaces, varying in terms of type of proprioceptive feedback. The performance in the turn task indicates what information is actually used for spatial updating when only path integration is being used.

2.2.1 Method

Participants

Ten participants from different universities took part in the experiment – eight males and two females – all with no prior experience in VE. The participants ranged in age from 19 to 26 years and had normal or corrected to normal eyesight. They all gave written consent and were paid a fixed amount for their cooperation. A financial bonus was offered to stimulate the participants to complete all the sessions with low variance.

Task

To determine their accuracy of path integration, the participants were asked in each condition to realise a series of instructed turning angles of 45Ñ, 90Ñ, 180Ñ, or 270Ñ in both clockwise and counter-clockwise directions and as accurately as possible (no time limit was given). As explained in the introduction, the angles of 90Ñ, 180Ñ, and 270Ñ were chosen to correspond to the accurate orthogonal reference frame that people have. The angle of 45Ñ was added to see whether a bias existed for angles other than the reference angles. Note that the 45Ñ angle is still larger than the FOV that is seen in the HMD (24Ñ), so that no direct overview of the angle as a whole is present.

Although participants were asked to perform the movement in one fluent motion (as opposed to moving in multiple blocks), they were allowed to and indeed encouraged to fine-tune their orientation at the end of each movement to correct for their perceived under or overshoot.

Participants were explicitly instructed not to move their heads in relation to their torsos while turning, because doing that would have provided them with additional kinesthetic feedback.

Design

A repeated-measures design was used with the within-participants factor *navigation interface*. In the two standing conditions (III and V), a total of 4 (angles) X 2 (directions) X 2 (measurements for each treatment) = 16 trials were carried out. With the other three conditions, this total was multiplied by 3 for the three different gains used for the input-device thus giving a total of 3x16=48 trials for these conditions. The order of the five conditions was balanced using a 5x5 Latin square, in which every sequence was completed by two participants each (Appendix A1). Within each condition the different instructed angles and mouse-gains were administered randomly. To prevent the wires from turning too much, the random generator for the angles was programmed so that the total angle turned over all trials did not exceed two whole turns.

Stimulus Conditions

Five navigation interfaces were tested, three with and two without visual feedback, in which three different methods were used to steer the rotation (Fig. 2.1).

I) In the purely *visual condition (Vis),* no vestibular and kinesthetic feedback from the body could be used to produce the instructed angle. However, there is still some extra-retinal feedback information available, generated by eye movements, which cannot be excluded. Participants were seated and used an elastic SpaceMouse™, of which only the rotational degree-of-freedom was enabled, to turn around in the virtual forest.

To prevent participants from using the proprioceptive feedback from the hand steering the input-device, a rate control was implemented with three alternative values for the gain. By changing the gain randomly after each trial, the proprioceptive feedback from the hand has no predictable relation with the turning speed. The gain values were chosen so that the maximum possible turning speeds were 45Ñ/s, 67Ñ/s, and 90Ñ/s. These speeds correspond roughly to the range of speed a person uses to turn using his or her legs while standing. The different gains also discouraged participants from using the unnatural strategy of counting time to approximate the turned angle.

II) In the *visual + vestibular condition (Vis + Vst),* participants could use vestibular feedback while turning in addition to the same visual feedback as in Condition I. No kinesthetic feedback from muscles of the legs could be used because the participants were seated on a large-sized turntable powered by an electro-motor. Their binaural axis was positioned approximately straight above the rotation axis of the turntable. The turntable speed was controlled by the participants using the input-device in the same way and with the same random gains as in the previous condition. The image of the virtual forest was kept in the correct perspective by head tracking.

III) In the *visual + vestibular + kinesthetic condition (Vis + Vst + Kin)* participants were standing and used their legs to turn around their body, thereby getting proprioceptive feedback both from the vestibular and from the kinesthetic senses as well as visual feedback. No input-device was used since the participants controlled turning with their body motion.

IV) The blind *vestibular condition (Vst)* is identical to Condition II, without visual feedback.

V) The blind *vestibular + kinesthetic condition (Vst + Kin)* is identical to Condition III, without visual feedback.

The same visual stimulus which was used for all three conditions with visual feedback was a virtual forest with 400 trees and 300 bushes randomly placed within a 500m radius circle, displayed on a stereoscopic HMD. All the trees and all the bushes were identical so they could not provide clear landmarks. Because configurations of trees and bushes could be recognised when making a full turn, the maximum turn angle chosen was 270Ñ. Since no unique visually recognisable cues were present, only the visual flow could be used to provide information about the angle turned. The depth information due to stereopsis is not essential for the perception of speed from visual flow. However, since it has been shown that depth increases the sensation of self-motion (Brandt, Wist, & Dichans, 1975), stereoscopic images were provided.

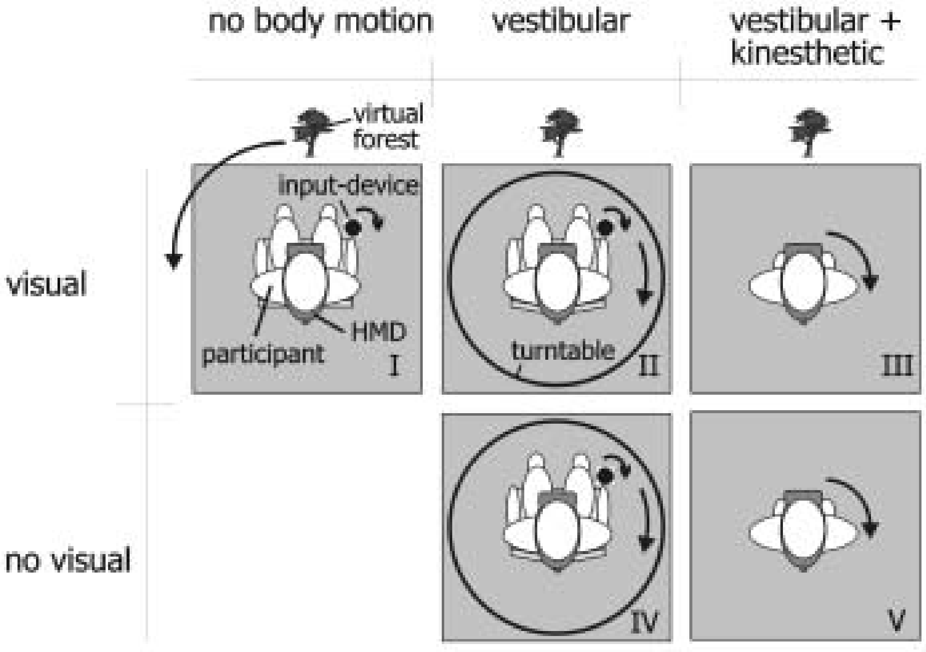


Figure 2.1 Navigation interfaces seen from above. For an explanation of the conditions (I, II, III, IV, and V) see the text.

A number of precautions were taken to prevent spatial orientation based on any cues other than those deliberately provided by the visual and proprioceptive stimuli described above:

- During all the trials, participants were given the same acoustic noise in their earphones. The spectrum of this noise was specially devised to mask sounds from the turntable motor and other specific sound sources in the laboratory room, like ventilator and computer noise. The required angles, generated beforehand by a random generator on a PC, were read to the participants before each trial by the experimenter on the same earphones using a microphone.
- To prevent participants from looking underneath the HMD, they were given a cap of thick black cloth, shielding them from all light from the environment.
- A completely new random forest was generated for each trial, so that no knowledge of the environment could be acquired, on which people could subsequently orient.
- To prevent participants from learning, absolutely no feedback about their performance was given after the trials or conditions. All trials for each condition were made in an uninterrupted sequence. Participants were therefore not able to see how much they had turned in between trials.

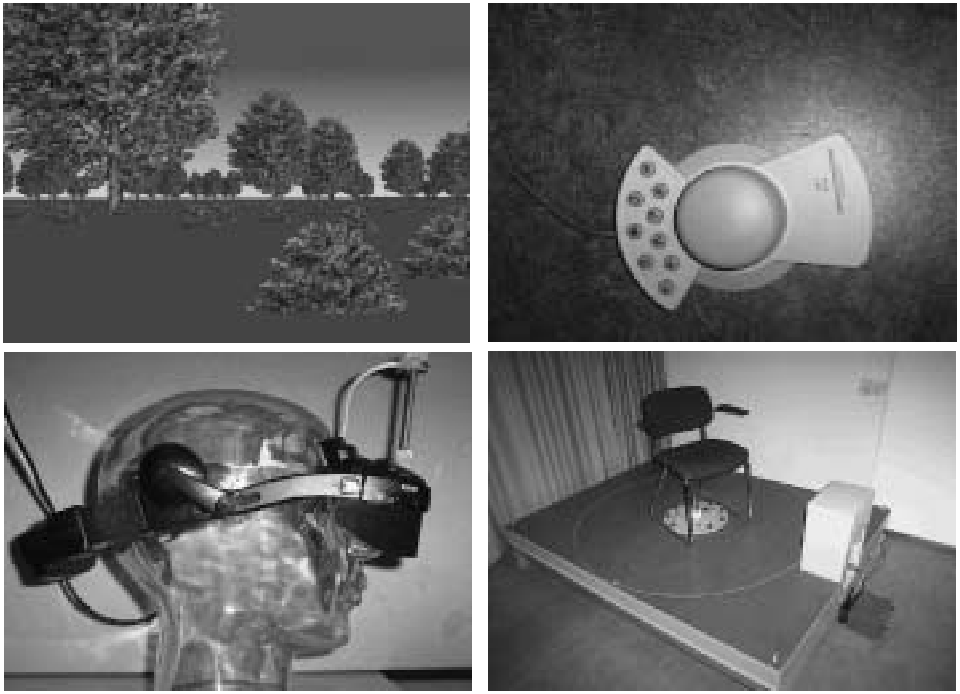


Figure 2.2 Apparatus and stimulus material. Top-left: image of the Virtual Forest (original stimulus material was in colour); top-right: the SpaceMouse input-device; bottom-left: the Virtual IO head mounted display with polhemus head tracking sensor; bottum-right: the turntable that was used for the vestibular conditions without kinesthetic feedback.

To enable turning, all wires from the HMD, the tracker, the SpaceMouse™ and the button were attached to the ceiling right above participants.

Apparatus

The images were generated by an Onyx Reality Engine with a Multi-Channel-Option that was manufactured by Silicon Graphics. The refresh as well as update rate was 60 Hz. The VE was drawn in linear perspective from a virtual eye point two meters above ground level. No shadowing was used because the shadows cast by a light source could provide unwanted directional cues.

A Virtual I/O stereoscopic HMD was used with a FOV of 24Ñx18Ñ (horizontal x vertical), a binocular overlap of nearly 100%, and an eye disparity that was 7.0 cm (Fig. 2.2).

Head position and orientation were tracked using the Fastrak™ (Polhemus). This electromagnetic tracking system consists of a transmitter that was attached to the ceiling, and a receiver that was attached to the HMD. Standard deviation in orientation measurements is less than one degree with a transmitter receiver distance of one meter or less (Werkhoven & Hoekstra, 1994). The delays between head movement and the corresponding visual feedback were approximately 80 ms.

The SpaceMouse™ from Virtual Technologies (Fig. 2.2) is an elastic six degree-of-freedom input-device with a knob approximately the size of an ice-hockey puck for controlling all motions. On the far side of the knob, there are nine buttons that can be used to provide discrete input. The resolution for each degree-of-freedom is 720 units for the full range of input.

The turntable (Fig. 2.2) is a horizontal disk of 1.5 m in diameter that can turn around a central axis powered by an electro-motor. A chair is mounted on top of the disk, which can be fixed in different positions relative to the centre. To control the turntable speed, the SpaceMouse™ input generated by the participants was fed into a program running on a 486DX/60 Mhz personal computer. A D/A converter was used to convert the output of this program to the analog signal needed to control the turntable engine. To keep the motions of the turntable sufficiently smooth, a limit was set for the acceleration of the turntable in the software prior to the experiment using a trial-and-error procedure.

Procedure

Prior to the experiment, participants were given a written explanation on the goals and the set-up of the experiment, after which there was an opportunity to ask questions.

Participants were given eight test runs before the actual trials started, in which they were able to become acquainted with the different methods for turning and with the different gains implemented for the input-device. As in the rest of the experiment, absolutely no feedback on task performance was given to the participants.

Participants were given a handheld button that they pressed to start and to finish each trial. The viewing direction of participants was registered at the beginning and at the end of each trial, together with the time of these events.

Scoring

The angle turned was calculated from the two angles registered at the start and finish of each trial. *Error* is defined as the angle turned minus the instructed angle. Positive errors therefore indicate that participants have turned too far (overshoot); negative error indicates that the participant has not turned far enough (undershoot). The *average speed* during a trial is calculated by dividing the angle turned by the duration of the trial. Note that undershoot means that participants overestimate the angle that they actually turn.

After each trial, participants were asked to rate their *subjective confidence* on a five-point scale in order to indicate how accurate they thought their produced angle was.

In between conditions, participants paused, filled in a questionnaire on *motion sickness*, and were interviewed by the experimenter about the strategies used to realise the required angles.

At the end of all five sessions, the participants were asked to rank the five navigation interfaces in a *ranking of difficulty* from one (easiest) to five (most difficult).

2.2.2 Results

One participant clearly showed deviant behaviour from the rest of the group. Since the results produced by this participant differed more than two standard deviations from the mean of the whole group, this participant was regarded as an outlier and was removed from the data-set, leaving a total of nine participants for analysis.

A post-hoc Tukey test for unequal N ($p<0.05$) showed no significant differences between clockwise or counter-clockwise turning errors. Therefore, the results of both turning directions were combined.

Error

Figure 2.3 shows the errors in the instructed turning angles for the five different navigation interfaces, averaged across directions, gains and participants. Participants tended to undershoot the instructed angle, with increasing undershoot for larger angles.

A two-way ANOVA *navigation interface* x *instructed angle* showed a highly significant main effect of *instructed angle* on error ($F_{3,24}=35.20$, $p<0.001$), and of *navigation interface* on error ($F_{4,32}=5.67$, $p<0.01$). The two-way interaction between *instructed angle* and *navigation interface* was also highly significant ($F_{12,96}=4.06$, $p<0.001$). To give an indication of the variation between participants, the standard deviations corresponding to the error means of Figure 2.3 are summarised in Table 2.1.

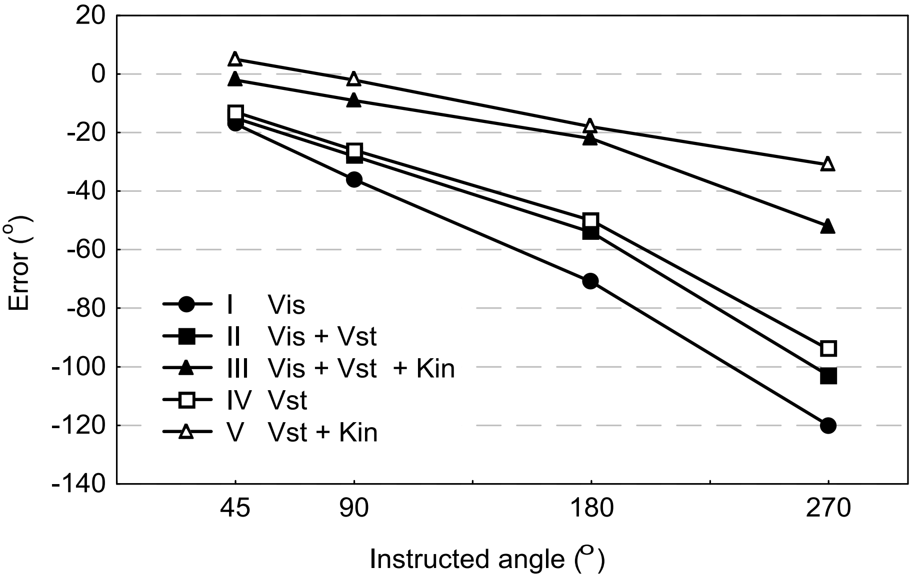


Figure 2.3 Mean errors for all five navigation interfaces (I to V). Positive values indicate overshoot; negative values indicate undershoot. Vis=visual; Vst=vestibular; Kin=kinesthetic.

Table 2.1: Standard deviations (Ñ) corresponding to the mean errors given in Figure 2.3. For each participant in each specific treatment condition (condition x instructed angle) a mean and a standard deviation was calculated for the four (III and V), or twelve (I, II and IV) repetition measurements (2 repetitions x 2 directions x 3 mouse-gains). The left part of the table gives the standard deviation between participants of the calculated means. The right part of the table gives the calculated standard deviation in the repetition measurements averaged over participants.

	Between participant standard deviations of the means				Mean of the within-participant standard deviations			
Condition	45	90	180	270	45	90	180	270
I Vis	17	32	51	77	8	12	22	27
II Vis + Vst	9	17	31	46	6	10	17	21
III Vis + Vst + Kin	6	9	25	29	5	9	15	18
IV Vst	8	14	39	51	7	12	21	30
V Vst + Kin	5	9	18	24	6	10	9	26

Looking at these standard deviations, an increase in variation is found for increasing instructed angles. It should be noted that the between-participant differences are lowest in the two kinesthetic conditions, followed by the two vestibular conditions, and highest in the purely visual condition. The within-participant variation differences between conditions are less substantial than the between-participant variation.

To show the effect of adding a visual stimulus to the vestibular or proprioceptive stimulus on error, Conditions II, III, IV, and V were analysed using a two-way ANOVA visual stimulus (present or not) and vestibular stimulus (with or without kinesthetic stimulus). Adding a visual stimulus or not, did not make a difference ($F_{1,8}$=1.9, p=0.2) and no interaction effect was found. The difference between vestibular and vestibular + kinesthetic was highly significant ($F_{1,8}$=22.8, p=0.001).

Input-device gain

The average turning speeds, given in Table 2.2, differ only marginally between conditions. The two standing conditions show the highest speeds followed by the blind vestibular condition. Since no time limit was given for the trials, participants sometimes waited longer than at other times before actually starting to turn. Since this waiting time is

Table 2.2: Average Speeds for Experiment 1. For Conditions I, II, and IV, speeds are given for the three different input-device gains separated by a '/'. Standard deviations correspond to variation between participants.

		Average speed	
Condition		Means (Ñ/s)	Std.Dev. (Ñ/s)
I	Vis	8/9/10	4/5/5
II	Vis + Vst	10/11/11	5/6/6
III	Vis + Vst + Kin	14	7
IV	Vst	11/13/13	6/7/8
V	Vst + Kin	21	10

included in the average speed, the actual turning speeds will be substantially higher. The three different input-device gains did not alter the actual average speed significantly. No significant main effect of the factor *gain* on error or on subjective confidence was found either.

Subjective confidence

Figure 2.4 shows the subjective confidence score of the instructed turning angles for the five different navigation interfaces.

Using a two-way ANOVA on *navigation interface* x *instructed angle*, highly significant main effects were found on subjective confidence of *navigation interface* ($F_{4,32}=11.3$, $p<0.001$), of *instructed angle* ($F_{3,24}=38.7$, $p<0.001$), and of the interaction between *navigation interface* and *instructed angle* ($F_{12,96}=2.01$, $p<0.05$). The general correspondence between the subjective confidence and the (objective) error score indicates that participants had a fairly good sense of the difference in accuracy between the different turned angles.

Ranking of difficulty

Figure 2.5 gives the results from the ranking of the difficulty for the navigation interfaces averaged over all participants.

Generally, the visual condition was judged to be most difficult, followed by the two turntable conditions. The two standing conditions (III and V) were found to be easiest.

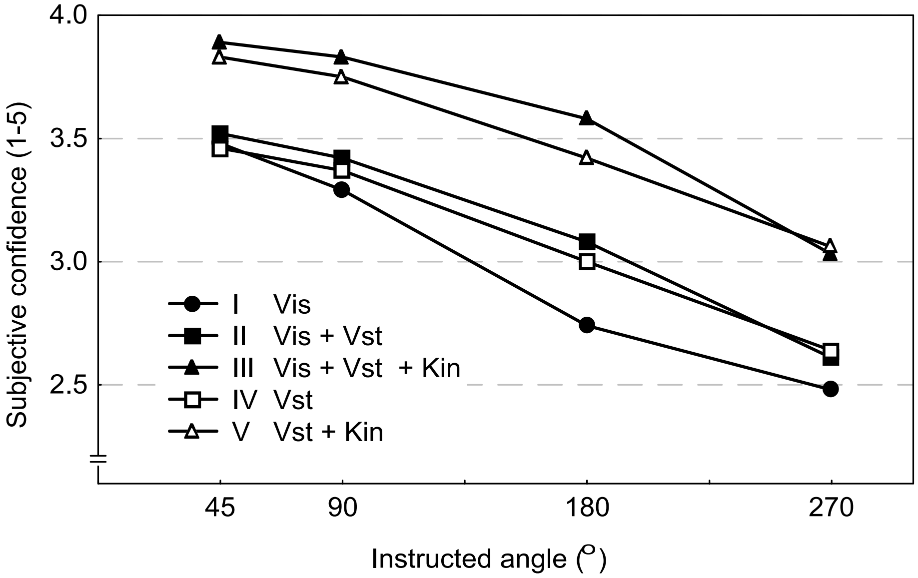


Figure 2.4 Mean subjective confidence for all five navigation interfaces (1 represents lowest confidence, 5 represents highest confidence). Vis=visual; Vst=vestibular; Kin=kinesthetic.

Analysing the ranking of difficulty with a non-parametric Friedman ANOVA showed a highly significant effect of navigation interface on ranking ($\chi^2_{N=9, df=4}=24.2$, $p<0.001$).

Interviews

From the interviews information was gathered on the strategies used for maintaining orientation. Several participants reported that they kept in mind the direction of the starting point relative to their body while turning. Other participants reported keeping in mind the direction of the destination point relative to their body. Some participants visualised the starting or destination point as a tree and others reported visualising a recognition point in the experimental room, like for instance the door (note that, since the experimental room could not be seen, this visualisation was purely imaginary). One participant reported imagining a circle divided into four parts with him turning in the centre.

Five out of nine participants reported that their attention was solely focused on their feet in the visual + kinesthetic + vestibular condition (III). The other four participants reported primarily using their feet, but also paid attention to the visual stimulus.

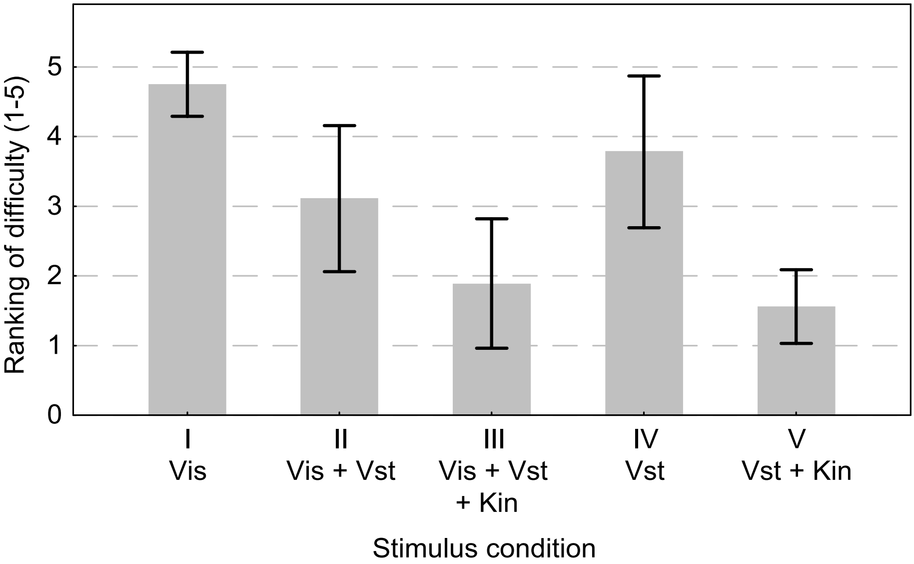


Figure 2.5 Participants' ranking of difficulty of the five different navigation interfaces (means and standard deviation); 1 represents easiest, 5 represents most difficult. Vis=visual; Vst=vestibular; Kin=kinesthetic.

In the questionnaires, participants did not report being aware of any systematic under- or overshoots. This was to be expected because if they had been aware of a systematic error, they would have corrected it.

Motion sickness

Finally, only one of the participants experienced a slight motion illness in the two standing conditions, but this did not, however, prevent him from finishing the experiment. Note that the duration of the sessions varied from approximately 6 to 10 minutes for the standing conditions, and 20 to 30 minutes for the other conditions. This exposure should be long enough to reveal participants' tendency towards motion sickness in the conditions experienced (Kolasinski et al., 1995).

Several participants reported getting a slight headache, probably from the HMD being pressed against their heads. Almost all the participants reported feeling stuffy as a result of the black cloth covering their heads.

2.2.3 Verification of positioning accuracy

In principle, it is possible that difficulties in accurate positioning with the different interfaces have contributed substantially to the errors found in the experiment. To verify whether this is the case, participants were required to orient themselves in the direction of a visual target in order to measure their positioning accuracy. Eight participants were used, all of whom were tested with the three different navigation interfaces that included a visual stimulus. Overall, a mean positioning error was found of -0.7Ñ, with a standard deviation of 1.4Ñ. Since this is substantially smaller than the errors found in Experiment 1, the positioning accuracy using the different navigation interfaces does not contribute to the errors found in the experiments.

2.2.4 Discussion

The best information for path integration, both objectively and subjectively, is provided by kinesthetic feedback. The errors for the kinesthetic condition in Experiment 1 are of the same magnitude as those that were found by Klatzky et al. (1990) and by Sadalla and Montello (1989). In the kinesthetic condition, participants do not have to estimate the turning velocity, instead they can estimate the angle covered by one footstep, and next they can count these steps to obtain the angle turned directly. Indeed, observations during the experiment showed that most participants tried to make 90Ñ steps to facilitate this process. In all other conditions, the turn velocity or acceleration had to be estimated from the stimulus and then integrated over time to obtain the angle turned.

Visual feedback in Experiment 1 apparently provides very poor information for path integration. The between-participant variation in the visual condition is substantially larger than in all other conditions. The performance in the purely visual condition was the lowest of all the conditions, although it was not significantly different from the two vestibular conditions. Besides, the addition of the visual stimulus did not significantly alter the results of the kinesthetic and vestibular conditions. This is reflected by the subjective results, indicating that, on average, the task in the visual condition is ranked the hardest. These results confirm the expectation based on Klatzky et al. (1998) and on Iwata and Yoshida (1999), that a virtual visual stimulus alone might not be sufficient to allow accurate path integration.

Vestibular feedback, although slightly better than visual feedback, does not enable the high performance level that is present if kinesthetic feedback is provided. The undershoot errors that were found in the vestibular condition (up to around 35 % at 270Ñ) are considerably larger than the undershoots (around 5 %) found by Israël et al. (1995). The difference might be caused by differences in acceleration magnitudes, but we have no way of verifying this.

In the visual and vestibular conditions, participants keep their average turning velocity constant, despite the different input-device gains. This indicates that participants do indeed estimate velocity from the given stimuli. However, looking at the errors given in Figure 2.3, the estimated speeds were obviously incorrect. The errors seemed to be approximately proportional to the instructed angle. This progressive undershoot can be explained as an overestimation of the turning velocity with a constant factor. The between-participant variation in the data (Table 2.1) shows that this velocity overestimation factor is not the same for different participants. The difference between participants in perceived velocity is smallest for the kinesthetic conditions followed by the two vestibular conditions, and largest for the visual condition.

A possible explanation for this misperception of speed in the visual condition (overestimation with a factor of approximately 1.7), can be found in the work of Wist et al. (1975). They found that the perceived speed of rotary self-motion increases linearly with the increasing perceived distance of the surroundings in a rotating drum. This is caused by a perceptual mechanism that serves velocity perception during transverse translations. In such cases perceived velocity is proportional to the retinal velocity of stationary objects multiplied by the perceived distance of the surrounding objects. Although this is correct for transverse translations, it is incorrect for rotations, where the rotation speed is only proportional to the retinal velocity and has no relation to the perceived distance of objects. Wist and his colleagues find that the perceptual mechanism for translating movement is also applied erroneously by participants to velocity estimation in rotating movements. Therefore, the rotation velocity is incorrect with a constant factor that depends on the perceived distance to the stimulus.

2.3 Experiment 2: Isolating the influence of visual feedback

From the results of the first experiment, two yet unsolved problems arise. Firstly, Experiment 1 showed that the addition of a visual stimulus to a kinesthetic or a vestibular stimulus does not significantly alter the turn performance of these conditions. However, it seems very unlikely that the presence of kinesthetic or vestibular feedback will totally rule out any visual stimulus influence. Secondly, it was assumed that participants extract velocity information in the purely visual condition from the visual flow and integrate this to realise the instructed angle. In principle, however, it is possible that participants just counted the number of trees passing a cross-section of the display, to measure the angle turned.

In Experiment 2, these two issues are explored by manipulating the visual flow, while keeping the number of trees passing a cross-section of the display per second the same as in the first experiment. This is implemented by changing the ratio of the physical FOV of the HMD and the displayed geometric FOV of the VE to 60%. Altering this ratio corresponds to introducing a zoom of 60%, which can be compared to looking through the wrong end of a pair of binoculars. The result is that the flow in Experiment 2 is reduced to 60% of the flow in Experiment 1, while speeds remain identical. The number of trees that pass a cross-section of the display at a specific speed, however, remains the same as in Experiment 1.

Firstly, The manipulation of the flow will clarify whether there is a visual stimulus influence when kinesthetic and vestibular feedback are present. Secondly, the manipulation will clarify whether participants just count trees, or whether they use the visual flow as information to execute the instructed angles. On the one hand, if participants do count trees, the same results will be expected as in Experiment 1, because the number of trees passing remains unaltered. On the other hand, if participants are using visual flow, they will be expected to overshoot the instructed angles by a factor 1/0.60 = 1.67 in this zoomed condition, assuming that their perception of the flow is precise. This would result in overshoot errors of 30Ñ, 60Ñ, 120Ñ, and 180Ñ, respectively, at the instructed angles of 45Ñ, 90Ñ, 180Ñ, and 270Ñ.

2.3.1 Method

Participants
Twelve students from several different universities participated – six males and six females – ranging in age from 21 to 27 years. All participants did not participate in the first experiment and they had no prior experience with VE. Participants had normal uncorrected eyesight. They gave written consent and were paid a fixed amount for their participation.

Task
The task was the same as in Experiment 1.

Design
A repeated-measures within-participants design was used (Appendix A.2). In the standing condition (VIII), a total of 4 (angles) X 2 (directions) X 4 (measurements for each treatment) = 32 trials per participant were carried out. In the other two conditions, this total was multiplied by 3 for the three different gains used for the input-device, giving a total of 3x32=96 trials per participant for these conditions. The order of the three conditions was counterbalanced, assigning two participants to each condition sequence.

Stimulus condition
As mentioned above, the second experiment differs from the first experiment in that a 60% reduction of the visual flow was created, at identical speeds. Note that the zoom artificially enlarges the geometric FOV in the VE to (24Ñx 18Ñ)/0.60 = 40Ñx 30Ñ. Since the same HMD was used as in Experiment 1, the zoom results in a compression of the image.

The user thus has an increased overview of the virtual forest. However, participants do not benefit from this overview, since there are still no recognisable features in the virtual forest, and the overview is still smaller than the smallest instructed angle of 45Ñ.

Since only the visual stimulus was changed in the second experiment, the two conditions without visual feedback were dropped, leaving three conditions for experiment two (see Conditions I, II, and III under Experiment 1, for a detailed description). To avoid confusion with the conditions of Experiment 1, numbering starts at VI and the z (from zoomed) is added to distinguish between the two different types of visual stimuli:

VI) visual (z)
VII) visual (z) + vestibular
VIII) visual (z) + vestibular + kinesthetic

Note that the vestibular and kinesthetic stimuli were not changed in the second experiment. The visual stimulus no longer corresponded to these proprioceptive stimuli as it would in the real world. In Experiment 1, these different stimuli did correspond to each other.

Apparatus and procedure

The apparatus and procedure are the same as in Experiment 1.

Scoring

As in Experiment 1, *error* is defined as the angle turned minus the instructed angle. The use of this definition is, however, not completely justified in the context of the purely visual condition of Experiment 2. The choice of the word “error” namely implies a deviation from a theoretically predicted angle (assuming perfect perception of velocity from the offered visual flow). In the purely visual condition of Experiment 1 the instructed angle corresponds to the theoretically expected angle. In Experiment 2, however, the visual flow is reduced by a factor 1.67 without the participants being aware of this. So, the theoretically expected angle is 1.67 times as large as the instructed angle, resulting in an overshoot error of 67% according to the definition of error as given in Experiment 1.

Note that the discussion above does not fully apply to the other two navigation interfaces in Experiment 2 involving proprioceptive feedback. For the proprioceptive feedback the instructed angle is the same as the theoretically expected angle, but a conflict is present between the scaled visual stimulus and the non-scaled proprioceptive stimuli.

To avoid confusion, the same definition of error will be maintained as that given in Experiment 1. However, the reader should note that in the purely visual condition of Experiment 2 an overshoot error of less than 67% of the instructed angle still means that the participant has not turned far enough and is overestimating the turning velocity from the visual flow point of view.

2.3.2 Results

Of the twelve participants, three were excluded from analysis which left nine participants. One participant was excluded from analysis because he was not able to keep his head upright during the trials. Two more participants were excluded from analysis because they reported using a visual strategy that overruled any influence of proprioceptive information in all conditions. These participants assumed that the FOV inside the HMD corresponded to a normal FOV without an HMD of approximately 180Ñ. To realise a turn angle of 90Ñ, for instance, they would calculate that a tree would have to traverse half of the display area.

The results produced by these participants showed similar results for all the navigation interfaces indicating that only the visual strategy was used, regardless of the added proprioceptive feedback. Besides, the turning speeds of the two excluded participants were substantially lower than those of the rest of the participants, indicating that their strategy might not be possible at high speeds. Note that the visual strategy mentioned here is not the same as counting the number of trees that passes a cross-section of the display. Although it would have been best to repeat the measurements with three new participants, this was not possible because the turntable was no longer available.

Error

Figure 2.6 shows the errors at the instructed turning angles for the three different navigation interfaces, averaged across directions, gains and participants. As was expected from the decreased visual flow, substantial overshoot angles were found in the purely visual condition. The overshoot is still substantially smaller than the theoretically expected 67% of the instructed angle (see the section on scoring in Paragraph 2.3.1).

A two-way ANOVA *navigation interface* x *instructed angle* shows a highly significant interaction between *navigation interface* and *instructed angle* ($F_{6,48}$=8.3, $p<0.001$). The main effect of *navigation interface* on error is also highly significant ($F_{2,16}$ = 7.8, $p<0.01$). However, the effect of *instructed angle* is not significant since the overshoot in the visual condition averages out with the undershoot in the other two conditions.

If one takes a closer look at the data using a post-hoc Tukey test ($p<0.05$), ones sees that there is a significant difference between the visual condition and the other two conditions with additional vestibular and kinesthetic feedback. However, no significant difference was found between the two stimulus vestibular feedback conditions with or without kinesthetic feedback.

To indicate the differences between participants, the standard deviations between-participants are summarised in Table 2.3. Looking at these standard deviations we see an increase in standard deviation for increasing instructed angle as in Experiment 1. Again,

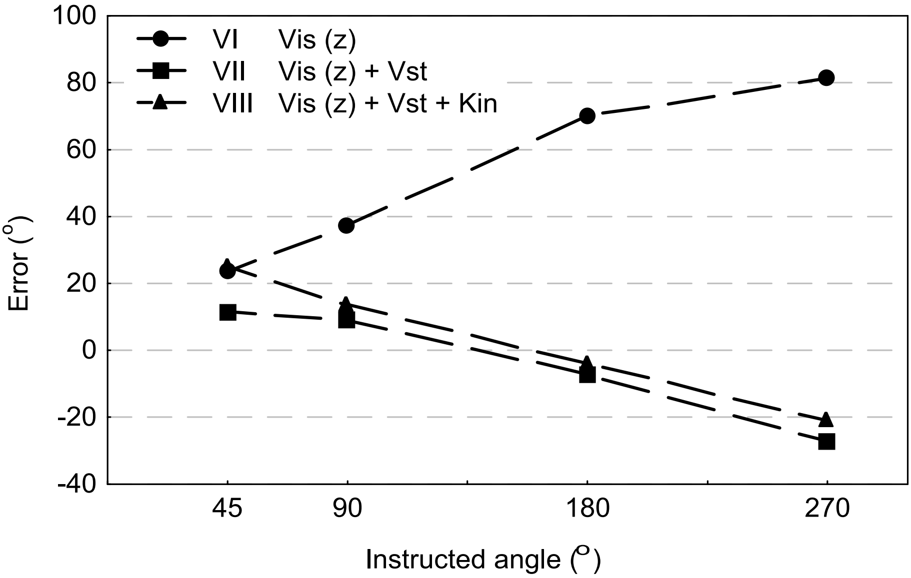


Figure 2.6: Mean errors for all three navigation interfaces (VI to VIII). Positive values indicate overshoot; negative values indicate undershoot. Vis=visual; Vst=vestibular; Kin=kinesthetic; z=zoomed.

Table 2.3: Standard deviations (Ñ) corresponding to the mean errors as given in Figure 2.6. For each participant in each specific treatment condition (navigation interface x instructed angle) a mean and a standard deviation was calculated for the eight (VIII), or twenty-four (VI and VII) repetition measurements (4 repetition x 2 direction x 3 mouse-gain). The left part of the table gives the standard deviation between participants of the calculated means. The right part of the table gives the calculated standard deviation in the repetition measurements averaged over participants.

	Between participant standard deviations of the means				Mean of the within-participant standard deviations			
Condition	45	90	180	270	45	90	180	270
VI Vis	23	30	61	88	19	27	52	64
VII Vis + Vst	16	24	37	56	12	17	29	36
VIII Vis + Vst + Kin	17	14	29	40	14	15	19	24

the between-participant variation is highest in the purely visual condition followed by the vestibular condition, and lowest for the kinesthetic condition. The within-participant variation follows the same pattern as the between-participant variation.

The turning speeds show clearly different results for the three different input-device gains (Table 2.4): higher gains result in higher average speeds and a larger variation.

Input-device gain

In contrast to the results of the first experiment, a two-way ANOVA *navigation interface* x *input-device gain* showed that the effect of *input-device gain* on error was significant ($F_{2,16} = 5.9$, $p=0.012$). A post-hoc Tukey for unequal N ($p<0.05$) reveals that the effect of *input-device gain* was only due to the visual condition, showing progressively increasing errors for increasing gain (errors of 38Ñ, 55Ñ, and 67Ñ, respectively). In the visual + vestibular condition, however, no significant effect of *input-device gain* could be found, although the turning speeds increase as the gain increases.

Subjective confidence

Figure 2.7 shows the subjective confidence score for the instructed turning angles and for the three different navigation interfaces. A two-way ANOVA *navigation interface* x *instructed angle* shows no interaction between *navigation interface* and *instructed angle*. The main effect on subjective confidence of *instructed angle* is highly significant ($F_{3,24} = 19.2$, $p<0.001$), and the main effect of *navigation interface* is only marginally significant ($F_{2,16} = 3.5$, $p=0.056$).

Table 2.4: Average Speeds for Experiment 2. For Conditions VI and VII, speeds are given for the three different input-device gains separated by a '/'. Standard Deviations correspond to variation between participants.

	Average Speed	
Condition	Means (Ñ/s)	Std.Dev. (Ñ/s)
VI Vis (z)	22/28/33	4/6/8
VII Vis (z) + Vst	21/25/26	4/5/7
VIII Vis (z) + Vst + Kin	21	6

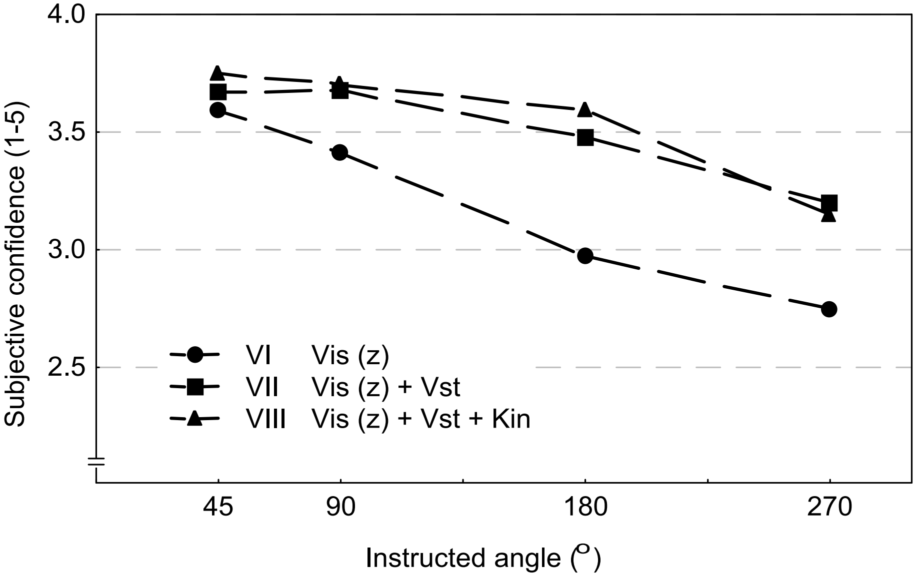


Figure 2.7: Mean subjective confidence for all three navigation interfaces (1 represents lowest confidence, 5 represents highest confidence). Vis=visual; Vst=vestibular; Kin=kinesthetic; z=zoomed.

A post-hoc Tukey test shows no significant difference ($p<0.05$) in subjective confidence between the two navigation interfaces with vestibular feedback (VII and VIII). The subjective confidence in the purely visual condition (VI) is, however, significantly lower than the other two conditions for angles larger than 90Ñ.

Motion sickness

Finally, none of the participants experienced motion illness. As in Experiment 1, several participants reported getting a slight headache that was probably caused by the HMD pressing against their heads. Again, almost all participants reported feeling stuffy as a result of the black cloth covering their heads.

Comparison between Experiment 1 and Experiment 2

In Experiment 1, participants had to turn prescribed angles being subjected to 5 different navigation interfaces. In Experiment 2, a zoom of 60% was introduced to the visual stimulus. To facilitate comparison between both experiments, the results of the two experiments are plotted together in Figure 2.8.

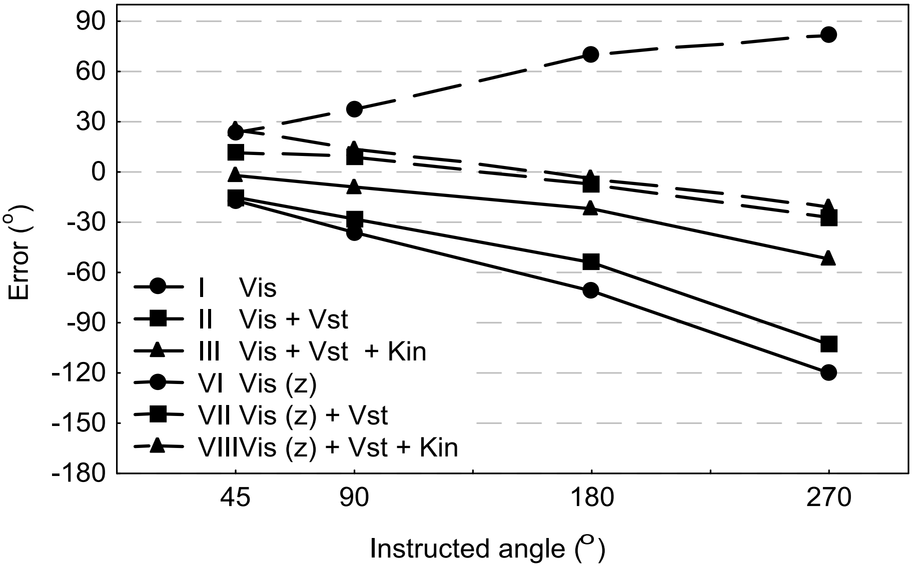


Figure 2.8: Mean errors for all conditions in Experiments 1 and Experiment 2. Positive values indicate overshoot; negative values indicate undershoot. Vis=visual; Vst=vestibular; Kin=kinesthetic; z=zoomed.

A t-test for independent samples was used to check whether there was a difference between having normal visual flow (Experiment 1) and having reduced visual flow (Experiment 2). Significant differences were found in all cases: for the two purely visual conditions (I and VI) t(16)=4.9; p<0.01; for the two turntable conditions (II and VII) t(16)=3.4, p<0.01; and for the two standing conditions (III and VIII) t(16)=2.8, p<0.05.

2.3.3 Discussion

In contrast to the undershoot errors found in Experiment 1, a substantial overshoot was found in Experiment 2 for the purely visual condition (up to 81Ñ error at 270Ñ, or 30 %). This overshoot is substantially smaller though than the 67 % overshoot that could be expected theoretically from the decreased visual flow (see the scoring section in paragraph 2.3.1). Therefore, participants in Experiment 2 were again not turning far enough and they overestimated their turning velocity, in a way consistent with Experiment 1. By contrast to the results of this study, Péruch et al. (1997) found that manipulating the geometric FOV in a triangle completion task had no effect. However, as was discussed in the introduction, visual recognition may have played a dominant role in their experiments.

The results provide an answer to the two questions that remained open after Experiment 1. Firstly, comparison between both experiments shows that the turn performance is significantly influenced by adding a conflicting visual stimulus to a proprioceptive condition. This means that the visual stimulus is not totally disregarded as was suggested by the results of Experiment 1. Secondly, the results show that participants do not use a strategy of counting trees. If participants had counted trees, the same undershoot would have been found in Experiment 2 as in Experiment 1. An alternative possible strategy is estimating the FOV inside the HMD as a reference angle, and calculating how many times a tree has to be followed from one side of the display to the other, in order to determine the angle turned. This strategy was indeed used by two participants who were excluded from the results, as mentioned before.

The combined proprioceptive + visual (zoomed) conditions lie in between the purely visual and the purely proprioceptive conditions. This seems to indicate that perception is based on a weighed average over the different available stimuli. Our results suggest that scaling the visual flow somewhere around 85% would actually compensate for the flaws in perception and lead to a near faultless mean turn production. As can be seen from the shift in lines in Figure 2.8, the conflicting visual stimulus of Experiment 2 seems to have a stronger effect on the turntable condition (vestibular) than on the standing condition (vestibular + kinesthetic). This corresponds to the fact that most participants report being primarily focused on their feet in the standing condition thus disregarding the visual stimulus.

Unlike in Experiment 1, input-device gain had a significant effect on error in the visual condition of Experiment 2, with higher gains showing larger errors. This means that the turning accuracy of participants becomes dependent on the configuration or gain that is chosen for the input-device.

The average turning speed in Experiment 2 was considerably higher than in Experiment 1, especially in the purely visual condition (28Ñ/s versus 9Ñ/s). It seems that participants turned faster in Experiment 2 to compensate for the decreased visual flow. Since participants controlled their turning speed themselves and no time pressure was given, we must assume that they optimised their speeds to facilitate the integration of the perceived stimuli.

No severe symptoms of motion sickness were found, whereas numerous cue conflicts were present. The absence of sickness is probably due to the fact that the task involved only a very specific motion pattern of movements around a vertical axis.

2.4 General discussion

Although the results found by Bles (1981) suggested that the use of different sensory modalities in the first experiment would lead to similar turn performance for all the different navigation interfaces, clearly distinctive results were found for these conditions. The purely visual condition resulted in the largest (absolute) errors, accompanied by the largest standard deviations, and the lowest subjective confidence. The vestibular conditions performed on the turntable, showed a modest improvement over the visual condition but did not yield the high performance that is found if kinesthetic feedback is present.

Participants generally tended to undershoot the instructed angles, which can be interpreted as an overestimation of the turning velocity. This is in line with the findings of Bles, Vianney De Jong, and De Wit (1984), who found systematic undershoots for angles of 360Ñ when participants have to step around in the dark. Similar undershoot angles were found when participants were allowed to complete the third leg of a triangle on the basis of only visual feedback (Péruch et al., 1997), or only on proprioceptive feedback (Loomis et al., 1993). Sholl (1989) found similar undershoot angles in 82% of the cases, when participants have to estimate the angle on the basis of vestibular feedback, being turned around in a wheelchair. Maybe this tendency to stop before a target is reached has been brought about by evolution to stop us in time from having dangerous collisions with objects or from falling into pits. The decrease in confidence and the increase in variation as the length of path integration increases, may lead us to make increasingly larger undershoot errors to maintain adequate safety margins.

The decrease in confidence and the increase in variation for increasing instructed angles in all conditions indicates that participants find it hard to integrate the perceived stimuli for long periods of time (or for long pathways). At the large angles, participants even incidentally reported, in the conditions without kinesthetic feedback, that they had totally lost track of the angle that they had turned.

In the experiments, participants were free to control their turning speed to facilitate their path integration ability. The fact that participants kept their average speed constant despite the different gains for the input-device shows that the turning speed is indeed important for path integration. Input-devices where the participants cannot control their speed freely may possibly reduce path integration performance.

Although this was not tested with the experiments, the increased overview, offered by the zoom factor of 60%, might be a strong aid to visual recognition. However, the experiments warn that path integration on the basis of only visual flow information will be strongly affected by manipulating the zoom. Furthermore, the performance in this case depends on the configuration (gain) of the input-device used.

The objective results showed a high level of correspondence to the subjective scores indicating the high validity of subjective evaluation methods for navigation interfaces as used by, for instance, Ware and Osborne (1990). The results of this study show that a larger absolute error resulted in lower subjective confidence. The ranking of the difficulty of conditions indicated clearly that participants preferred having kinesthetic feedback, followed by vestibular feedback, and finally visual feedback.

Apparently, using the body to move around in a VE is superior to using an indirect input-device to control movement. The kinesthetic feedback from the legs and feet provides direct turn angle information. This is clearly different from the information perceived in the other conditions of the experiments, where participants have to estimate the turning velocity or acceleration from the stimulus and integrate this information over time to obtain the turned angle. Darken and Sibert (1996) report on a participant who kept one foot fixed to provide an absolute reference to enable a systematic search pattern in a novel VE. This shows that an absolute kinesthetic reference can be very valuable to allow for certain kinds of exploration strategies for virtual worlds. The results of this study quantify the accuracy of such a reference.

2.5 Conclusions

The results have shown that proprioceptive feedback, particularly kinesthetic, can be used quite effectively for path integration. The perception of orientation from visual flow or vestibular feedback alone is inaccurate. The considerable accumulation of errors in path integration over time indicates that, even with kinesthetic feedback, visual recognition is still needed to ensure sufficiently accurate spatial updating.

The results show that path integration is enhanced by using special input-devices that provide natural kinesthetic feedback for rotations. If an input-device is used besides a head-tracking system in an immersive VE system, then directional control should not, if possible, be allowed by the input-device.

3

Training visual path integration[1]

Abstract

The experiments in Chapter 2 have shown that path integration is inaccurate if a non-immersive interface is used that provides only visual feedback, unlike an immersive interface that provides both vestibular and kinesthetic feedback.

In Experiment 3, what is explored is whether explicit feedback during training can compensate for the disadvantage of a non-immersive interface. Results show that by providing participants with Knowledge of Results (KR), they can indeed calibrate the biases in their path integration process, and they can also maintain their improved level of performance during a retention test carried out the next day.

Experiment 4 provides a verification of the visual stimulus by manipulating the FOV, the update rate, and the size and density of the stimulus. The stimulus verification showed a strong effect of changing the FOV, thus providing some explanation for the differences between Experiment 1 and Experiment 3.

In Experiment 5, participants are trained without feedback but distinct objects are put in the VE, thereby allowing visual recognition to be used to calibrate path integration. Results showed that visual recognition can be adequately used by participants to bring their bias in path integration down to the same level as that of explicit feedback.

---

[1] Part of this chapter will be published as:

Bakker, N.H., Werkhoven, P.J., & Passenier, P.O. (in press). Calibrating visual path integration in VE. *Presence*.

3.1 Introduction

The results of Experiment 1 and Experiment 2 showed that human path integration ability is poor for virtual rotations if based only on visual feedback. Results showed a strong tendency to overestimate the turning velocity in the condition with only visual feedback leading to undershoots if participants were asked to turn a specific angle. Furthermore, a large variability in performance was found both within and between participants indicating the difficulty that participants have with processing the visual stimulus. These conclusions were confirmed by other results relating to visual path integration (Péruch et al., 1997). Klatzky et al. (1998) found that participants did not update their mental representation of their orientation at all in a triangle completion task if they were shown an animation. In all these experiments, however, participants received no Knowledge of Results (KR) during training or testing. KR can be provided by giving participants feedback about the errors they make in performing their task.

In contrast to these poor results for visual path integration, Van Veen and Riecke (1999) found only small turning angle errors in a visual triangle completion task, but they gave their participants an extensive training with KR before their measurements were made. Repeated exposure to virtual environments may help users to learn the relationship between perceived visual flow and actual displacement in the VE. The orthogonal angles of walls in rooms and hallways in a virtual building may help users to learn how much visual flow corresponds to turning an angle of, for instance, 90Ñ. There is evidence that spatial orientation improves with repeated exposure in a non-immersive VE system (Ruddle et al., 1997; and Ruddle et al., 1998). Calibration of the visual path integration process may have caused part of the improvement in these experiments, but no direct evidence is available.

Although the results of Experiment 1 and Experiment 2 suggest that an immersive interface should be used to improve path integration, this might be undesirable for financial reasons or because of other unsolved negative side effects of an immersive interface, for instance the high occurrence of motion sickness. Providing a short training can perhaps relieve the limitations of visual path integration. To our knowledge, no experiments have been reported in the literature which directly investigate the effect of training with KR on path integration with a non-immersive navigation interface.

To see whether participants can actually learn what is the relation between visual flow and the angle turned, the effect that receiving KR during training has on visual path integration is investigated. Part of Experiment 1 on path integration was repeated, using the condition with just visual feedback, but now one of two groups of participants were supplied with KR. By providing KR, it is expected that participants can learn the relationship between visual flow and the actual displacement. It is thus expected that the large errors in turning an instructed angle, as found in the earlier experiment, will be reduced.

3.2 Experiment 3: Explicit training by provision of Knowledge of Results (KR)

3.2.1 Method

Participants

Thirty-eight colleagues and students, 19 male and 19 female, served as participants. The participants ranged in age from 19 to 31 years and had no prior experience with VE. Twenty of the participants were used in the KR group and eighteen participants formed the control group. The males and females were divided approximately equally over both groups. It was necessary to have a larger number of participants assigned to the KR group to increase reliability for the retention test, which was not carried out by the control group.

Task

Participants were instructed to turn angles in the virtual forest *as accurately as possible* without any time limit. The instructed angles were 90Ñ, 180Ñ, or 270Ñ, in both left and right directions. Although the participants in our experiment were asked to perform the movement in one fluent motion, as opposed to stopping on the way, they were allowed to and indeed encouraged to fine-tune their orientation at the end of each movement, to correct their perceived under or overshoot. Participants controlled their turn rate in the virtual forest by turning the elastic knob of an input-device.

Design

The participants were randomly assigned either to the KR group or to the control group. The KR group had to perform four sessions: a *pre-test* session, without feedback on performance; a *test* session, with KR feedback; and two retention sessions, again without feedback. The first retention session was held five minutes after the test session, and the second retention session was held a full day after the test session. The control group performed only the first two sessions, but both without KR. The control group was included in the experiment to verify that any improvements found in the training of the KR group would only be due to the provision of KR.

The first and last two sessions consisted of 36 trials and the second session consisted of 54 trials. The test session consisted of a larger number of trials to give participants more opportunity to improve their performance. During the test session, the instructed angles were grouped first as four, then as three and then as two trials, so participants in the KR group could immediately use the feedback from the previous trial to try to improve their performance during the next trial. In all other sessions the order of all the different angles was random.

Stimulus conditions

The visual stimulus consisted of a virtual forest with 400 trees and 300 bushes randomly placed within a 500m radius circle, and displayed on a stereoscopic HMD. All the trees and the bushes were identical both size-wise and in appearance, so they could not serve as landmarks for visual recognition. A completely new random forest was generated for each trial, to avoid allowing the participants to learn and use the configurations of trees and bushes.

The gain of the input-device used to turn around was chosen randomly for each trial from three alternative values. This was done to ensure that the proprioceptive feedback from the hand alone could not provide participants with information regarding turning speed. Maximum turning speeds with the three gains were 80Ñ/s, 160Ñ/s, and 240Ñ/s. To speed up task execution the gains were larger than in Experiment 1 and Experiment 2. The different gains made it impossible for participants to rely only on counting time to correctly estimate the angle turned. Furthermore, since cues for visual recognition were absent, only the integration of the visual flow could be used to provide information about the angle turned.

Apparatus

Participants were seated in a chair with a wooden board on their lap, on which the input-device was attached to controll the turning speed. The input-device was the SpaceMouse™ (Virtual Technologies), where only one degree-of-freedom was used. A number of buttons is present on the input-device, which the participant used to indicate the start and finish of each trial.

Images were generated by an Onyx Reality Engine with Multi Channel Option manufactured by Silicon Graphics. A Kaiser ProView60 stereoscopic HMD was used with a physical FOV of 48Ñ (H) x 36Ñ (V), a binocular overlap of 100%, and adjustable eye-pieces. The geometric FOV was set equal to the physical FOV of the HMD and the eye disparity was set fixed to 7.0 cm. To prevent participants from being distracted by light coming from underneath the HMD, the room in which the experiment took place was darkened. For more details about the apparatus see Chapter 2.

Procedure

Prior to the experiment, an explanation was given to the participants about the goal and the set-up of the experiment, after which there was an opportunity for them to ask questions. Before starting the sessions each day, participants were given four trials to familiarise themselves with the trial procedure.

Participants started each trial by pressing a button, after which the instructed angle was shown in the upper right part of the display and the participants could start turning. When the participants thought they had realised the instructed angle they indicated this by pressing the button again, after which the next trial could be immediately started.

During the test session, the participants in the KR group received feedback about the errors that they had made in the right upper part of the screen, immediately after each trial. A '+' was shown if errors were below 10Ñ. If larger errors were made, the feedback consisted of one, two, or three arrows pointing in the direction where the participant should have been. The number of arrows indicated the magnitude of the error: one arrow means 10Ñ< *error*Ä 30Ñ; two arrows means 30Ñ<*error*Ä 60Ñ; and three arrows means *error*>60Ñ. The arrows were chosen to avoid confusion with the signed values of instructed angle.

Scoring

Absolute error (Ñ), that is to say the absolute value of the signed error, is taken as the main performance measure to see whether performance improves with KR provision. To ascertain whether changes in performance are due to a change in participants' bias or to a change in variation of performance, *signed error* and *within-participants variation* are also taken as dependent measures. *Signed error* (Ñ) is defined as the difference between the realised angle and the instructed angle. The realised angle is the difference between the angles registered at the start and finish of each trial. Positive values indicate overshoot and negative values indicate undershoot. *Within-participant variation* for a single participant in a specific group at a specific angle is the standard deviation of the measurements that are repeated. For Sessions 1, 3 and 4 with 36 trials each, this means that the standard deviation is taken at each angle over 12 trials. For Session 2 with 54 trials, the standard deviation is taken at each angle over 18 trials.

The *Average speed* of turning is calculated to register the possible changes due to the different mouse-gains. To get a good estimation of the average turning speed, a timer was triggered when participants had turned the first 15Ñ, ensuring that the time participants waited to actually start turning was not included in the calculation of the average speed. The *average speed* during a trial was calculated, therefore, by dividing the total angle turned (including movement for corrections) minus the start angle (15Ñ) by the duration of the trial from the moment of passing the start angle. Because the first turning part is left out, the average speed can not be used as an absolute indication of the speed, but only for relative comparison between mouse-gains.

3.2.2 Results

Since no differences could be found between left and right turning, the results of both directions were combined. The results produced by one participant in the control group were excluded from analyses because this participant had misinterpreted the task. This participant reported that he imagined himself moving around the perimeter of a large circle while looking outward, instead of turning on the spot as both the instruction and the stimulus material indicated. The errors made by this participant were sometimes over a thousand degrees, far greater than the errors of all other participants.

The effect of KR

To investigate the effect of feedback, the absolute error data of the first two sessions was analysed with a three-way ANOVA with the (between participants) factor *treatment group* (2 levels), and the (within participants) factors *session* (2 levels), and *instructed angle* (3 levels) (Fig. 3.1, and Table 3.1).

The effectiveness of training is apparent from the reduced absolute errors for the KR group in Session 2 where feedback is provided (Fig. 3.1). The significant interaction between *treatment group* and *session* (TS in Table 3.1) indicating the effectiveness of training, is analysed in detail with a post-hoc Tukey test for unequal N ($p<0.05$). This shows that performance is significantly improved for the KR group (from 69Ñ in Session 1, to 23Ñ in Session 2) whereas no significant increase in performance was found between

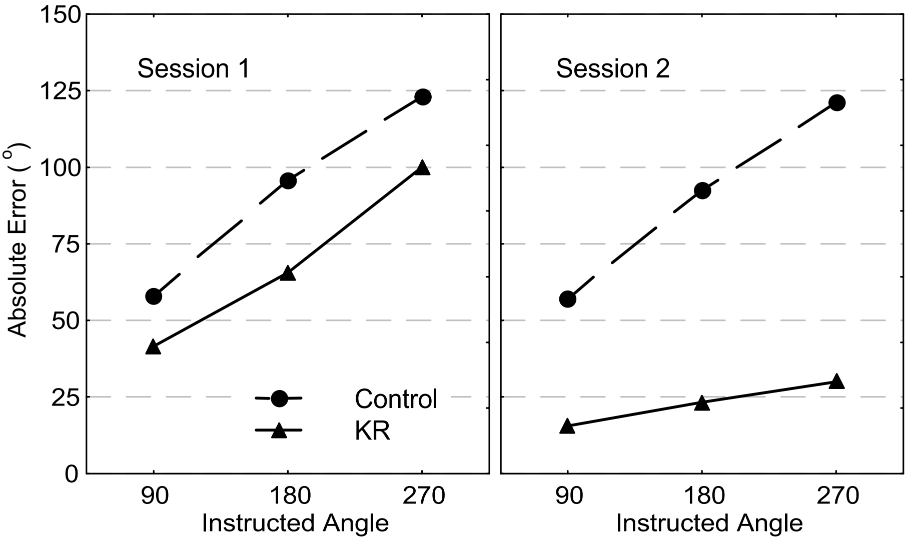


Figure 3.1: Three-way interaction of *treatment group* x *session* x *instructed angle* for absolute error (Ñ) averaged over participants, mouse gains, and repetitions. KR indicates Knowledge of Results.

Session 1 and 2 (92Ñ to 90Ñ) for the control group. Furthermore, the difference in absolute error between the KR group and the control group in Session 1 (69Ñ versus 92Ñ) was not significant.

Apart from the positive effect of KR, the errors increased with larger instructed angles. The significant three-way interaction indicates that the increase in absolute error with larger instructed angles is reduced for the KR group in Session 2 where KR is provided. A post-hoc Tukey test ($p<0.05$) shows that there are no significant differences between the three instructed angles for the KR group in Session 2.

Retention of training

The absolute errors for all sessions are given in Fig. 3.2. The retention of training for the KR group is tested by comparing the absolute errors for all sessions, using a post-hoc Tukey test ($p<0.05$). Performance improvement for the KR group was retained if KR was removed in Session 3 and also when testing the next day in Session 4 occured.

There were large differences between participants in the means of absolute error (Fig. 3.2). These differences were largely reduced for the KR group in Session 2 where KR was provided, and remained small for this group in later sessions.

Table 3.1: Main effects and interaction of a three-way ANOVA on treatment group (T) x session (S) x instructed angle (I) for Session 1 and 2.

	df	F	p
T	1,35	8.29	<0.01
S	1,35	13.72	<0.001
I	2,70	50.32	<0.001
T x S	1,35	11.49	<0.01
T x I	1,70	4.20	<0.05
S x I	2,70	8.26	<0.001
T x S x I	2,70	7.87	<0.001

Bias

To get an indication of whether the improvement found with KR is due to a reduction of bias, the signed error is examined. Signed errors for the control group were 13.0Ñ and 17.8Ñ for Session 1 and 2, and for the KR group 35.3Ñ, –0.4Ñ, 5.6Ñ, and 13.1Ñ for Sessions 1 to 4. An analysis of Session 1 and 2 using a three-way ANOVA *treatment group* x *session* x *instructed angle* shows only a significant interaction between *session* and *treatment group* ($F_{1,35}$=4.83, p<0.05). Detailed analysis of this interaction with a post-hoc Tukey test (p<0.05) shows that the bias is significantly reduced for the KR group and no reduction is apparent with the control group. Looking at data produced by only the KR group for all sessions with a two-way ANOVA with factors *session x instructed angle*, only a non-significant trend was found for *session* ($F_{3,57}$=2.55, p=0.06). The large number of effects found with the absolute error are not found with the signed error due to the cancellation of negative and positive values.

Variation

To get an indication whether the improvement found with KR is due to a variation reduction for individual participants, the within-participant variation was examined. An analysis of Session 1 and 2 with a three-way ANOVA *treatment group* x *session* x *instructed angle* shows only significant main effects for *session* ($F_{1,35}$=12.21, p<0.01) and for *instructed angle* ($F_{2,70}$=41.16, p<0.001). Mean within-participant variation for Session 1 and 2 averaged over the instructed angles are 42Ñ and 32Ñ, and for the instructed angles (90Ñ, 180Ñ, and 270Ñ) averaged over the first two sessions are 25Ñ, 39Ñ, and 48Ñ. Since

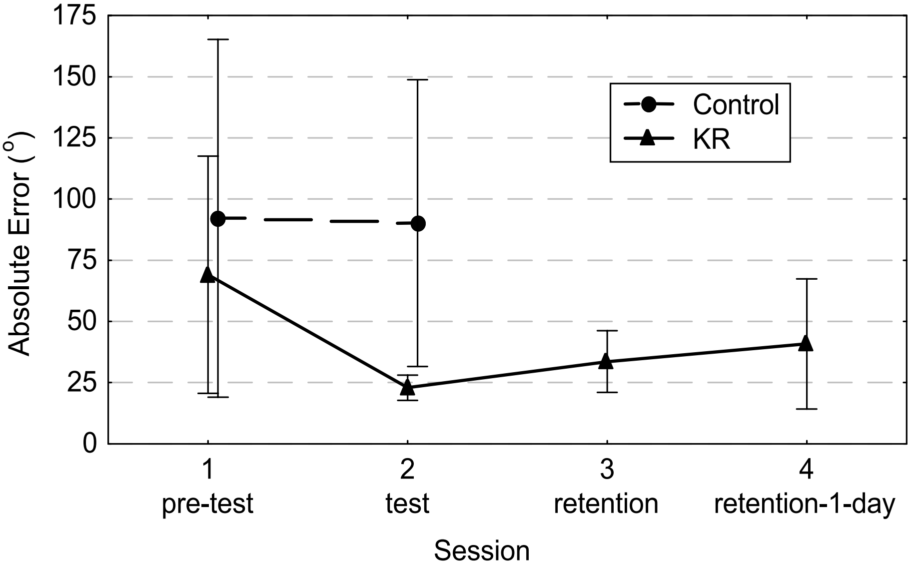


Figure 3.2: Absolute error (Ñ) averaged over instructed angles, participants, mouse gains, and repetitions; whiskers indicate one standard deviation corresponding to differences between participants. KR indicates Knowledge of Results.

there is no significant interaction, no advantage was found with the KR group above the control group in terms of a reduction of within-participant variation.

Looking at data for only the KR group for all sessions with a two-way ANOVA *session* x *instructed angle*, significant effects were found of *session* ($F_{3,57}$=14.15, p<0.001), of *instructed angle* ($F_{2,38}$=70.13, p<0.001) and of the interaction between the two (F6,114=4.29; p<0.006). The interaction shows a clear division between a strong effect of angle in Session 1 (26Ñ, 40Ñ, and 56Ñ) and a more moderate effect of instructed angle in the other three sessions (18Ñ, 26Ñ, and 34Ñ averaged over the last three sessions).

Effect of Mouse-gain on Error

The effect of Mouse-gain on absolute error was investigated for Session 1 and 2 with a three-way ANOVA *treatment group* x *session* x *mouse-gain*. Only the results of main effects and interactions that include the mouse-gain effect are discussed because the effects of the other factors on absolute error have already been reported above. Increasing the mouse-gain has a significant effect on absolute error (60Ñ, 70Ñ, and 75Ñ for increasing mouse-gains; $F_{2,70}$=8.65, p<0.001). A post-hoc Tukey test shows that the absolute errors with the two highest gains are no different. No interaction of the factor mouse-gain with the other factors was found, indicating no significant change in the effect of mouse-gain for different sessions or for the different treatment groups.

To see whether the effect of mouse-gain is caused by bias or by within-participant variation, three-way ANOVAs were also performed for the signed error and for the within-participant variation. Signed error increases significantly (5Ñ, 19Ñ, and 26Ñ as mouse-gains increase; $F_{2,70}=15.8$, $p<0.001$). A post-hoc Tukey test again shows no difference between the two highest gains. Within-participant variation also increases significantly (43Ñ, 46Ñ, and 51Ñ for increasing mouse-gains; $F_{2,70}=4.12$, $p<0.05$). A post-hoc Tukey test shows that the results of the middle gain do not differ from the results found in the lowest or the highest gain.

Effect of Mouse-Gain on Average Speed

The effect of Mouse-gain on average speed is investigated for Session 1 and 2 with a three-way ANOVA on *treatment group* x *session* x *mouse-gain*. Apart from a significant main effect for gain on speed (22Ñ/s, 31Ñ/s, and 38Ñ/s for increasing mouse-gains; $F_{2,70}=51.2$, $p<0.001$), a three-way interaction is found between all factors. Inspection of the data shows that this is caused by a reduction from Session 1 to Session 2 of the speed, but only for the highest two mouse-gains of the KR group (Session 1: 38Ñ/s and 41Ñ/s, and Session 2: 30Ñ/s, 35Ñ/s). Furthermore, a significant interaction was found between *session* and *treatment group* ($F_{1,35}=4.94$, $p<0.05$). The interaction is caused by a slight increase in speed between Session 1 and Session 2 for the control group (from 29Ñ/s to 30Ñ/s), and a slight reduction in speed between Session 1 and Session 2 for the KR group (from 33Ñ/s to 29Ñ/s). No other effects were found.

3.2.3 Discussion

As expected, a substantial reduction of absolute errors in path integration based on visual flow can be achieved by providing participants with Knowledge of Results (KR). The performance improvement remains present, even when testing a full day later. However, the trend in the data shows that the errors increase again after KR removal, suggesting that the retention of training might fail for a longer duration than the one day that was tested here. Judging from our results, the low errors found by Van Veen and Riecke (1999) in comparison with Péruch et al. (1997) in a triangle completion task, might be attributed to the KR that they provided for participants during training.

Our experiment only investigated the updating of heading. What remains to be investigated is whether the reduction of bias for turning that was shown here, also provides an advantage for more complex tasks like learning a building layout (e.g. Ruddle et al., 1997; or Ruddle et al., 1998).

Further inspection of signed error and within-participant variation shows that the performance improvement due to KR can be especially attributed to a reduction in bias and not to a reduction in the variability of participants' performance. The variability is larger in Session 1 than in later sessions, but no differences between treatment groups were found. The higher variability in Session 1 was probably due to the fact that participants had only four trials for familiarisation before starting Session 1, so they were still trying out different strategies.

The instructed angle has a highly significant effect on absolute error. This is to be expected from an integration process because biases cause increasing error if length or duration of integration increases. The effect of instructed angle disappears if KR is provided in Session 2 to the KR group. This is a logical consequence of the reduction of the bias to near zero (-0.4Ñ) for this group in this session. Most results regarding the bias are largely obscured by the fact that some participants showed negative biases, whereas others showed positive biases that cancelled out against each other in the analysis.

Manipulation of the mouse-gain has a significant effect on performance, which is due to the larger biases as well as the increased variability for larger mouse-gains. However, the effect on absolute error is small (a difference of 15Ñ between largest and smallest gain) in comparison with the effects of instructed angle (differences over 60Ñ) or the effect of providing feedback to the KR group (difference is 46Ñ). Larger mouse-gains lead to a larger average turning speed of participants. When KR is provided, participants reduce their speed for the highest two gains.

The increase in speed (from 22Ñ/s for the lowest gain to 38Ñ/s for the highest gain) is too large in comparison with the increase in error to suspect that participants used a strategy of counting time instead of estimating the turning speed. It is suspected that participants had difficulty controlling their speed with the increased sensitivity at higher mouse-gains. However, since the velocity profiles were not recorded during the trials, there is no evidence to support this.

The large absolute errors found in both the control group and the KR group in Session 1 before actual training, correspond to the results of Experiment 1 and 2 where no KR was given. However, the signed errors that were found in Experiment 3 indicate that participants overshoot the instructed angle on average, whereas in the former study participants showed a tendency to undershoot the instructed angle. Because both studies use similar stimulus material, task, and procedures, this difference is probably caused by the different FOVs (24Ñ versus 48Ñ horizontally) that were used in the studies due to the different HMDs that were used (Virtual/IO versus Kaiser ProView60 in this study). The within-participant variation that was found here for the three instructed angles of the last three sessions (18Ñ, 26Ñ, and 34Ñ) are of the same order of magnitude as was found in our prior experiment (12Ñ, 22Ñ, and 27Ñ).

3.3 Experiment 4: Effects of FOV, update rate and virtual forest set-up

In Experiment 1 and Experiment 2 on the one hand and Experiment 3 on the other hand, similar but not completely identical stimulus material was used. A factor that might have contributed to the differences between the result is the different FOVs (24Ñ versus 48Ñ horizontally) that were used in both the studies. Furthermore, the tree texture was changed between Experiment 2 and Experiment 3 because the original texture was lost due to partly failing software backups. This might have influenced the apparent size and distance of the trees. Also, in Experiment 2 the zoom was manipulated, which meant that the relation between the geometric FOV and the physical FOV was changed. Besides changing the optic flow such a manipulation has two side effects. Firstly, the number of trees visible within the current view is increased. Secondly, the size of the trees is reduced, again influencing the apparent size and distance of the trees. Although no effect would be suspected from these factors verification is necessary.

A factor possibly influenced by the changes in trees is the update rate that might have been different between the experiments although this was not registered. Update rate is partly determined by the number of overlapping polygons in depth that is possibly increased by the larger tree size in Experiment 3.

3.3.1 Method

To investigate the effects of FOV, update rate, and virtual forest set-up, a control experiment was performed directly after Experiment 3 with the same participants. Because this control experiment was only thought of halfway through the experiment, only the last 23 participants were used (11 of the control group and 12 of the KR group). To limit the number of necessary trials, participants were only required to turn 180Ñ to the right. Each condition was repeated five times and the results of the repetitions were averaged for each participant. Five stimulus conditions were investigated:

1. Normal: the stimulus material is identical to the stimulus in Experiment 3.
2. FOV restricted: In this condition the FOV is restricted to 20Ñ horizontally by fixing a virtual black mask in front of the viewpoint, while using the same HMD. The relation between the geometrical FOV and the physical FOV is *not* changed by this manipulation. The virtual forest is identical to the first condition.
3. Number of trees halved. In this condition the number of trees is halved in comparison with Condition 1.
4. Size of trees doubled. In this condition the size of all trees and bushes is scaled by a factor of two.
5. Update rate limited: In this condition low and variable update rate is generated by including an extra delay of 0.1 seconds in the image generation loop once in every four to seven times. This does not affect the turning velocity that is determined on a separate computer by the input-device excursion. The virtual forest is identical to Condition 1.

Participants were told that the stimulus would be manipulated, but not precisely how.

Before and after the trials were administered, participants were interviewed regarding the strategy that they used to produce a turn and the difficulty of the five conditions. The preliminary interview concerned the trials performed just before Experiment 4 in Experiment 3, by the same participants.

3.3.2 Results

The data were analysed for *signed error*, w*ithin-participant variation* and *absolute error* with three separate two-way ANOVA *stimulus condition* x *treatment group*. There is a highly significant main effect of *stimulus condition* ($F_{4,84}$=26.33, p<0.0001) on signed error. A post-hoc Tukey test shows that this effect can solely be attributed to the different path integration behaviour with the restricted FOV (Fig. 3.3). No main effect was found of *treatment group* although Fig. 3.3 clearly indicates a trend in this direction, nor was there an interaction between *treatment group* and *stimulus condition*. There were no effects on absolute error, probably because the negative error found for the restricted FOV condition was cancelled. For within-participant variation, only a main effect was found of *treatment group* ($F_{1,21}$=5.07, p<0.05) indicating higher variability of performance for the control group (45Ñ) than for the KR group (24Ñ). This was already found in Experiment 3.

Interviews

The interview beforehand provided much insight into how participants performed their task in Experiment 3. Participants reported two different strategies when asked how they had performed the task. The first strategy used most by eleven participants, was that of picking a tree and following it to the edge of the display. After the tree disappeared

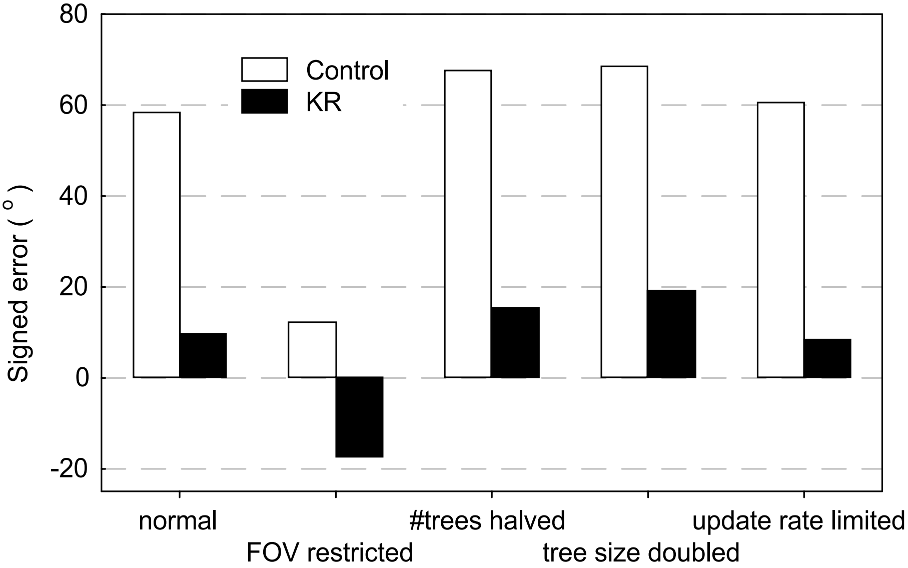


Figure 3.3: The effect of the display condition on signed error (Ñ). The means were averaged over all participants. The instructed angle was 180Ñ in all the trials. KR indicates Knowledge of Results.

beyond the edge of the display, it was followed mentally until it reached 90Ñ. The displacement of the tree outside the FOV must somehow be estimated from the visual stimulus. For angles of 180Ñ and 270Ñ the process was repeated twice or thrice. The second strategy, reported by six participants, involved trying to estimate the velocity and the duration of turning without consciously following specific trees. Two of these participants reported fixing their gaze on the instructed angle sign displayed in a fixed spot on the screen. The remaining six participants could not report the strategy that they had used. Four participants reported that they always finished with a tree in the centre of the display.

The interview following Experiment 4 partly confirmed the objective results obtained for the different conditions. When asked which stimulus condition was hardest, 16 participants said that the restricted FOV condition had been the most difficult one, 6 participants said that the limited update rate condition was the hardest one, and one participant found no differences in the level of difficulty. As second hardest the limited update rate was mentioned 15 times and the restricted FOV 4 times. Four of the six participants that found the limited update rate condition the hardest, had also reported using the strategy of estimating velocity.

Ten participants commented on turning with the enlarged trees. Nine of these spontaneously described the condition with the enlarged trees as the condition with the trees nearby. This shows that the size cue strongly affected the apparent distance of the trees. Five of the ten participants reported that they thought they were turning faster or had to turn for a shorter duration; three other participants found this condition easier than the normal condition; and two other participants found turning with the enlarged trees harder than the normal condition.

When asked to estimate the FOV in the VE after the HMD was removed participants responded divergently and inaccurately. The answers were between 45Ñ and 180Ñ, the average being 100Ñ. If participants had used this FOV estimation to calculate how many times they had to follow a tree from the middle to the edge of the display, the results would be completely different in Experiment 3 and Experiment 4. With this strategy, the overestimation of the FOV would result in undershoot errors, whereas in fact overshoot errors were found.

3.3.3 Discussion

The results of Experiment 4 suggest that the size of the FOV partly explains the differences in signed errors between Experiment 1 en Experiment 3. In Experiment 1, undershoot errors were found while using an HMD with a FOV of 24Ñ whereas in Experiment 3, overshoot errors were found in the same task while using an HMD with a FOV of 48Ñ. Experiment 4 showed that a reduction in the FOV makes participants turn considerably smaller angles (37Ñ less with reduced FOV) in concordance with the differences between Experiment 1 and Experiment 3.

Neither the size nor the number of trees had any effect on path integration performance, reconfirming that the manipulation of zoom in Experiment 2 did not affect performance by means of changes in tree size or tree density.

The limited update rate that was created by introducing 0.1 second delays in the image-rendering loop clearly disturbs the optic flow. A tree would be displaced over 24Ñ from one frame to the other during the delay, if a participant moved at the maximum speed of 240Ñ/s. In this case the impression of a smooth continuous motion is broken. The visual stimulus can only provide information for the participant about the rotation if the participant recognises the specific tree after displacement.

The fact that participants are not affected by the limited update rate suggests that they integrate displacement rather than velocity from the visual stimulus. The reports on the strategies used confirm that many participants consciously follow a specific tree at least during the first part of turning.

3.4 Experiment 5: Implicit training by visual recognition

In Experiment 3 it was shown that participants can adequately calibrate visual path integration for rotations if provided with KR. The explicit feedback was provided by indicating the error with a number of arrows. A less explicit form of training would be to put uniquely identifiable objects in the VE. Participants could then use visual recognition to calibrate path integration. Participants who face a specific object and make a full turn would encounter the same object again, thereby knowing the turn sensation corresponding to 360Ñ. If participants are not allowed to carry out such an inspection but can only make small turns from different starting points this would be harder because the information has to be integrated over several turns. To verify whether participants are able to use visual recognition to calibrate their visual path integration, a second control experiment was performed directly following the first, but now in a world with uniquely identifiable objects and only small turns.

3.4.1 Method

After the control experiment reported in the previous paragraph, the same 23 participants performed the 90Ñ turning task in an environment where clearly distinctive objects were placed. Participants were explicitly instructed not to look around to explore the world. They were required to restrict themselves to the turning trials. Two virtual worlds were constructed, both with enough objects so that there was always at least one unique object in view (Fig. 3.4). The objects were, for instance, a water tower, a barstool, a cow, a bench, or a car. In each of the two worlds, participants were instructed 16 times to turn 90Ñ to the right. After each turn the participant was put in different starting orientation in a systematic way. The starting orientations were 0Ñ, 180Ñ, 45Ñ, 225Ñ, 90Ñ, 270Ñ, 135Ñ, and 315Ñ after which the same series was repeated again. This series was chosen to minimise the chance of encountering the same orientation in two consecutive turns.

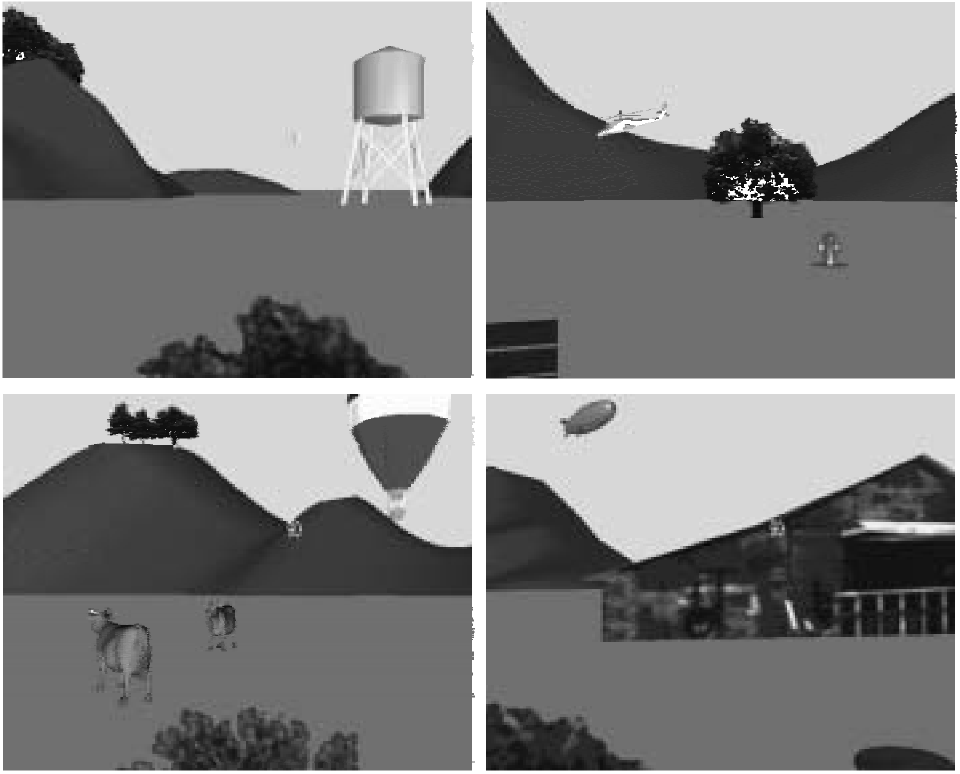


Figure 3.4: Four typical views of the one of the worlds with distinct objects (original stimulus material was in colour).

3.4.2 Results

The data were analysed for signed error, within-participant variation and absolute error with three separate two-way ANOVA on *stimulus condition* x *treatment group*.

The results for signed error and absolute error show the same effects. Significant interactions were found between *treatment group* and *session* ($F_{2,42}=6.01$, $p<0.005$ and $F_{2,42}=7.92$, $p<0.001$, respectively) and a significant main effect of *session* ($F_{2,42}=9.11$, $p<0.001$ and $F_{2,42}=4.60$, $p<0.02$, respectively). For the KR group no differences in performance were found between the last session in Experiment 3 and the two sessions in the world with objects. The control group improved performance dramatically when visual recognition was possible (Fig. 3.5). Signed errors decrease in later sessions for the control group (61Ñ, 11Ñ, and 9Ñ), but remain constant for the KR group (11Ñ, 4Ñ, and 8Ñ). The within participant variation shows a significant interaction between *treatment group* and

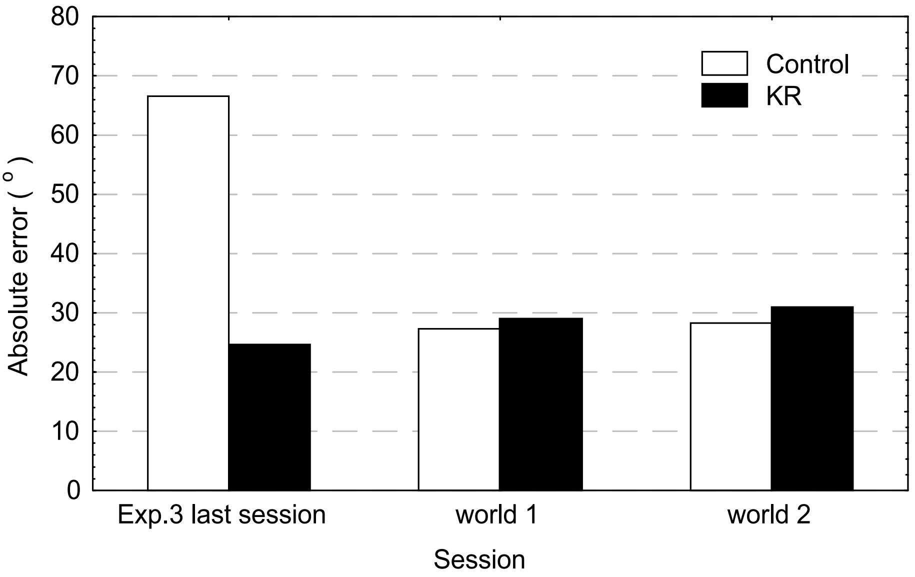


Figure 3.5: The interaction between session and treatment group on absolute error. The instructed angle was 90Ñ in all the trials. KR indicates Knowledge of Results.

session ($F_{2,42}=4.16$, $p<0.02$) and a main effect of *treatment group* ($F_{1,12}=5.44$, $p<0.03$). A post-hoc Tukey test shows that the within participant variation decreases for the control group (31Ñ, 25Ñ, and 19Ñ) and remains constant for the KR group (15Ñ, 19Ñ, and 19Ñ).

3.4.3 Discussion

The sudden performance improvement for the control group in a virtual world where visual recognition can be used is remarkable. Remember that participants were not allowed to look around in the world before starting to turn. The general tendency of participants in the control group to overshoot (signed error 61Ñ in Experiment 3) must have made it easier to detect the incorrectness of their turn responses because they simply get to see more of the VE which they can use to calibrate path integration. The improvement to the same level of the KR group suggests that explicit training with KR might not even be needed to calibrate path integration.

3.5 General discussion

The effectiveness of training.

It was expected that in providing participants with knowledge of results they would be able to calibrate their path integration process, learning the relation between visual flow and the actual displacement.

The results of the training showed that indeed the bias could easily be corrected. Even with less explicit feedback the biases disappear by allowing visual recognition of distinct objects in the VE to occur. The different path integration errors found in the literature, as mentioned in the introduction (Péruch et al., 1997; Klatzky et al., 1998; and Van Veen and Riecke 1999), might very well be explained by differences in feedback provided to participants during familiarisation with the experimental apparatus. The results of this study show that the biases in the visual path integration process can be calibrated at least temporarily by providing participants with KR. However, training visual path integration does not help to reduce variability of performance for individual participants.

Differences in bias between the experiments

Experiment 3 showed different biases in visual path integration from Experiment 1 in Chapter 2. In Experiment 1 participants had a strong tendency to undershoot the instructed angle, whereas in Experiment 3 overshoot errors were found. To find an explanation for these differences a control experiment was performed in which the visual stimulus material was varied so as to reflect the differences between the experiments.

The results showed that an explanation for the differences in bias between the experiments may be sought in the different HMDs used in both studies. Verification of the stimulus material showed that the bias is strongly affected by the physical FOV that is to say by the choice of specific HMD. A smaller physical FOV leads to smaller turned angles, corresponding with the direction of the differences that were found between the experiments.

The strong effect of FOV on path integration performance suggests that maybe other parameters of the optical set-up of the HMD could also have an influence. The distance between the left and right eye virtual viewpoints is held constant throughout the experiments whereas the inter pupillary distance varies from one participant to another. Besides, in the experiments reported in this chapter, an HMD was used with adjustable eyepieces, which may create small differences in the FOV for participants. There was no simple way of checking how participants positioned the displays in front of their eyes.

Turning strategy

Large differences were found between individual participants' biases in Experiment 3 but also in the earlier experiments in Chapter 2. Some of these differences may be caused by differences in the strategies that participants use. The results of Experiment 2 ruled out the possibility that participants might use the strategy of counting trees. Interviews showed that participants used two main strategies.

The first strategy used most by participants was that of registering their displacement by following a specific tree. This strategy is relatively insensitive to low update rate as long as there is enough overlap between two successive frames. The tree that is followed can then still be matched with the corresponding tree in the next frame after a large displacement between two frames. The correspondence problem might increase the difficulty of the turn task for a low frame rate as reported in the interviews, but this does not need to affect performance as with the objective results. Supporting these results, Vishton and Cutting (1995) found for translation that heading perception is determined by displacements and is relatively invariant when there are changes in frame-rate. Besides, they showed that the sensitivity to low update rate increased if finding correspondence over frames is made harder by making the visual objects less distinct. The velocity or the duration of turning is of no importance to this strategy. Since the eyes move to follow objects, the extra-retinal eye signals could, in principle, be used by the path integration process.

The second strategy that was used by participants involved estimating the turning velocity in combination with the duration of turning. Two participants reported fixing their gaze on the instructed angle sign that was displayed in a fixed location on the screen. This would mean keeping their eyes stationary and having no extra-retinal eye signals available.

Participants might very well use both strategies. In fact, many participants who used the displacement strategy reported that after following an initial tree to the side of the display, they would continue following this tree mentally until it reached 90Ñ. How they used the visual stimulus during this last part of the turn can only be speculated upon.

The results that were obtained on the the strategies employed were subjective reports. Participants might not be fully aware of the strategies that they use or may switch between strategies depending on the information that is available. To discriminate between the strategies objectively, experiments would have to be performed with eye-tracking equipment to register eye movements, and to perform more manipulations of the visual stimulus.

Variability of performance

Besides differences between participants there is also considerable within-participant variation. A part of the within-participant variation is caused by an effect of the random mouse gains. In Experiment 3 the difference in the signed errors between the lowest and highest mouse-gain is 21Ñ. In Experiment 1 no effect of mouse-gain was found on error. In Experiment 2 the difference between the lowest and the highest mouse-gain is 29Ñ. Another source of variability for some participants is the tendency to stop precisely with a tree in the centre, which was done by at least four participants. Increasing the density of the visual stimulus would reduce this variability. However, no effect was found of increasing the number of trees in Experiment 4.

When a tree is followed that crosses the edge of the FOV, the participant must move his eyes rapidly to the other side and select a new tree to follow. This transition moment may be an extra source of variation or even of systematic bias.

3.6 Conclusions

The biases in visual path integration are strongly affected by the physical FOV. The visual path integration process can easily be calibrated to reduce the existing biases. This calibration can be accomplished by providing knowledge of results or by putting participants in a VE with distinct objects that can be visually recognised. The performance improvement remains, but the trend in the data suggests that the biases will increase again in the long run. The variation in performance is not reduced by the training, which indicates that even after training, the visual component of the path integration process is unreliable, at least with a virtual visual stimulus.

4

Acquisition of a cognitive map[1]

Abstract

In two experiments Hypothesis 2 and Hypothesis 3 are explored. Immersive navigation is expected to improve the acquisition of a Cognitive Map (CM) when compared to non-immersive navigation. Furthermore, navigation by discontinuous displacement is expected to attenuate the acquisition of a cognitive map when compared to navigation by continuous movement.

Experiment 6 investigates how fast participants can acquire a CM of the layout of a VE if provided with either an immersive or with a non-immersive navigation interface and with either continuous or discontinuous displacement. After learning the VE the disorientation is measured when participants are displaced discontinuously to a random orientation. Results indicate that inexperienced participants learn the VE faster with continuous displacement than with discontinuous displacement. However, these differences disappear in the long run. Between the immersive and the non-immersive interface, no differences were found in learning duration. After being displaced to a random orientation participants are temporarily disoriented.

Experiment 7 investigates whether path-integration training beforehand improves the acquisition of a CM with non-immersive navigation. The difference between immersive and non-immersive navigation is investigated again. Results show that there are no benefits to be gained from path-integration training in the acquisition of a CM. Although no difference in learning duration was found, a test after learning shows a better quality of CM for the immersive than for the non-immersive navigation condition.

---

[1] Part of this chapter has been published as:
Bakker, N.H., Passenier, P.O., & Werkhoven P.J. (1998). Spatial orientation in virtual environments: isolating the roles of head-slaved vision and continuous visual feedback. *Proceedings of the 17th European Annual Conference on Human Decision Making and Manual Control, Valenciennes, France*, pp.187-196.

4.1 Introduction

The experiments in Chapter 2 and 3 have shown that navigation with an immersive interface leads to more accurate and reliable path integration than navigation with a non-immersive interface. In these experiments, visual recognition and cognitive anticipation could not be used for spatial updating. If all three spatial updating processes can be used, then the effect that differences in path integration quality may have on the internal representation of the current location will depend on how information from the three processes is integrated.

Furthermore, the information available for path integration depends on whether the navigation is continuous or discontinuous. With discontinuous displacement the feedback cues that are normally available for path integration are absent. Therefore, the relative contribution of the three processes to spatial updating depends on the type of displacement.

As was explained in Chapter 1, the greatest role of path integration in spatial updating is expected when exploring an unknown, novel environment (McNaughton, Knierim, & Wilson, 1995; O'Keefe & Nadel, 1978, p94; and Shenk, 1998). In this situation, a CM is not yet present, which implies that visual recognition of the own location is not yet possible. Once the environment has been learned, visual recognition and cognitive anticipation may become more dominant in spatial updating and the role of path integration may diminish.

Both the experiments that are reported in this chapter consist of two phases. In the first phase the acquisition of a CM is investigated under the influence of different navigation interfaces and different types of displacement. Participants have to learn the locations of objects placed on an imaginary cylinder surrounding them. The time that participants take to learn the locations is a measure of the efficiency of spatial knowledge acquisition under the different conditions.

In the second phase, after having learned the layout, a number of tests are performed to see if differences exist between the quality of the CMs built using the different navigation conditions. Furthermore, these tests serve as a measure of how much time participants take to recover from disorientation after discontinuous displacement. Participants have to search the objects that have been learned, starting either from an orientation that is known in advance or from a random orientation.

The first phase of the first experiment in this chapter, Experiment 6, investigates the effects on spatial knowledge acquisition of both the type of interface (immersive versus non-immersive) and of the type of displacement (continuous versus discontinuous).

With the continuous displacement condition the images were updated by 20Hz, providing a fluent visual movement. In the discontinuous displacement condition, the space that surrounds the participant is divided into non-overlapping adjacent sections (Fig. 4.1).

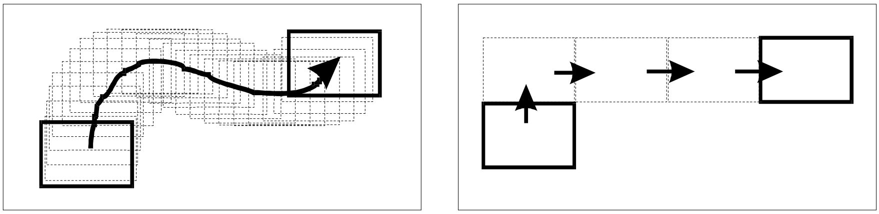


Figure 4.1: The continuous (left) and discrete (right) display conditions. In the continuous condition, the view moves smoothly through the VE. In the discrete condition, the view moves discretely to adjacent non-overlapping sections of the VE.

Interface type and displacement type were combined full factorially, leading to four conditions:

- In the non-immersive continuous condition, participants can use visual recognition as well as visual path integration to determine their location while exploring the VE.
- In the immersive continuous condition, participants can also use proprioceptive feedback for path integration, on top of the cues that are already available with the non-immersive continuous condition.
- In the non-immersive discontinuous condition, participants can no longer use path integration to determine their location. Participants only see static images. To relate the information that is seen in one image to the next image, knowledge is needed about the spatial separation between the viewpoints that correspond to the images. This knowledge is provided to participants beforehand in the instructions. In such cases cognitive anticipation can be used to relate the images.
- The immersive discontinuous condition is somewhat unusual. Participants use their body to control movement immersively. However, the visual feedback that they receive is discontinuous. If a participant moves around within a section, a frozen image is seen, but as soon as the border between two sections is crossed, the image switches instantaneously to the image corresponding to the next section.

If path integration is used in spatial updating, then learning the VE in the immersive conditions should be faster than in the non-immersive conditions, because path integration quality in this first condition is superior. The non-immersive discontinuous condition is expected to lead to the poorest learning performance result because path integration cannot be used, and on top of that, cognitive anticipation needs to be used to relate the images. The results for the immersive discontinuous condition are harder to predict. On the one hand, participants can use path integration based on proprioceptive feedback, which may support spatial updating. Participants can use the path integration information to determine in which section they are located. On the other hand, there is a discrepancy between the static image and the changing location that a participant feels while moving around within one section. Because there is no fixed relation between the current location as indicated by path integration and the location as observed by visual recognition, it might be hard to encode the object locations in the CM.

In the second phase of the experiment, the efficiency of spatial updating is investigated after two types of discontinuous displacement have taken place. With the first displacement type the participant is always put in the same start-orientation prior to having to find an object that participants learned in the first phase. Participants can use cognitive anticipation to reduce disorientation by the discontinuous displacement. With the second displacement type the participant is put in a random orientation. Participants can only use visual recognition to determine their location after the displacement.

As a measure of spatial updating efficiency, the time that participants take to anticipate to the displacement is registered, as well as the time that participants need after displacement to recognise their location. Furthermore, the number of times that participants choose a wrong route to the object that has to be found is registered.

If any difference in quality exists in the CMs acquired in the different conditions, these are expected to be reflected in the reduced efficiency of spatial updating in the tests. The time given to prepare for the fixed orientation includes the time needed to adapt to the new own location, the time to remember the target-object location, and the time to determine the shortest route to the target's location. With the random start orientation the time to prepare includes only the time needed to remember the target-object location. When compared to the random start orientation, the fixed start orientation is expected to show an increase in preparation time, but a decrease in the time needed for recognition as well as a decrease in the incorrect route choices.

The first part of the second experiment in this Chapter, Experiment 7, investigates whether path-integration training improves the acquisition of a CM with non-immersive navigation. Chapter 3 showed that the biases in path integration that are present for the non-immersive interface can easily be reduced by training. However, the high variability in visual path integration performance remains present, even with training. Whether this reduction of bias affects spatial updating if visual recognition can also be used is unknown. An experiment is performed, which is similar in set-up to Experiment 6, in which participants have to learn the location of objects that surround them, using different navigation interfaces.

Three conditions are tested, all with continuous displacement: a non-immersive navigation interface, a non-immersive navigation interface with path-integration training beforehand, and an immersive navigation interface that serves as a baseline. If poor quality of path integration with a non-immersive interface leads to an increase in the time required to learn an environment, then training of path integration beforehand is expected to ameliorate spatial knowledge acquisition.

After learning a spatial layout in one of the three conditions, participants are again tested on the finding of objects from different start orientations. On top of the test conditions that were already present in Experiment 6, a start orientation is included in which no displacement was made after the previous trial. The results of this *uninterrupted start orientation* are expected to be similar to the fixed start orientation results, with the exception of the preparation time which should be smaller for the uninterrupted start orientation. However, if cognitive anticipation to a discontinuous displacement is not able to compensate completely for the loss of path integration information then the uninterrupted start orientation should also lead to lower recognition times or less incorrect route choices.

4.2 Experiment 6: The effect of the navigation interface and of the type of navigation

4.2.1 Method

Participants

Sixteen participants, eight males and eight females, took part in the experiment. The participants, who came from different universities, ranged in age from 19 to 29 years, and had normal or corrected to normal eyesight. Participants gave written consent and were paid a fixed amount for their participation. A bonus was given to the eight best participants for their performance. Two participants did not complete the experiment and were replaced by two new participants to complete the design. The first of these two participants suffered from motion sickness after the discontinuous immersive condition and the second was not able to complete more than two sessions within four hours.

Task

For each participant the experiment consisted of four sessions, one for each of the four different stimulus conditions. Each session consisted of a learning phase and a test phase. In the learning phase, emphasis was put on learning and remembering the positions of all the objects **in the least number of trials possible** without time restraints. In the learning phase, the participants carried on until they had a good internal representation of the virtual world. The learning phase consisted of blocks of six trials in which all of the six different objects were presented once as the target objects to be searched for. In between blocks a written test with questions about object locations was performed to determine whether the knowledge was adequate and learning could be stopped.

In each trial, a participant had to find one of the six objects in the virtual world. The trial started by showing a small version of the object that had to be found in the upper right side of the display. This target object remained on this fixed screen position throughout the trial. The participants then had to push a start button to make the rest of the virtual world visible. Next, participants had to search until they had found the target object. In front of the target object a small display gave either the number 1 or the number 4, in correspondence with the mouse button which the participants had to press to end the trial. The number in front of the target object was changed randomly between the two options for each trial to ensure that participants would have to find the target object to successfully end the trial.

In the test phase, the participants again had to search for the objects, but this time they were instructed to take **the shortest possible route** to the target object and to perform the task **as fast as possible**. Participants were instructed to think first where the target was located, before pressing the start button. This allowed us to measure the time it took participants to retrieve the location of the object from their CM.

Design

The factors *navigation interface* (immersive versus non-immersive) and *displacement* (continuous versus discontinuous) were combined factorially, resulting in four different conditions, tested in four separate sessions. The four conditions were tested within-participants, and the order of conditions was counterbalanced using a 4x4 Latin Square, resulting in four different sequences (Appendix A3). In each session, participants had to learn the layout of six objects in a virtual world using the specific navigation interface for that session. Four different worlds were used, ensuring that the layout and the appearance of the objects were different for each session. Within each of the four condition sequences, the sequence of the four different worlds was again counterbalanced using a 4x4 Latin Square giving a total of 16 unique combinations of condition and world sequences. Four participants were tested for each sequence of conditions, and each was given a different sequence of virtual worlds.

Spatial knowledge acquisition was performed in blocks of six trials each. In between blocks, participants performed a virtual world knowledge test on paper to determine whether they had attained sufficient knowledge. After learning the VE, to finding of objects was tested using two different types of start-orientation: a *fixed* orientation that corresponded to the situation during learning, or a *random* orientation. Participants performed three fixed and three random blocks alternately comprising 12 trials each, and starting with a fixed block. Participants were told what treatment they would receive. The first two blocks, one fixed and one random, were regarded as training and were not analysed. In the random condition, participants first had to figure out how they were oriented in the virtual world, before they could decide on the shortest route to the target object.

Stimulus conditions

Navigation interfaces

With the two immersive conditions, the viewing direction was determined on the basis of the orientation measured by the tracker mounted on top of the HMD. Therefore, the body movements of the participant directly controlled the viewing direction, as is normally the case in immersive VE systems. Participants were seated on a swivel-chair and could use their feet to push the chair around. In the case of the two non-immersive conditions the heading and the pitch were controlled by an input-device. The other degrees-of-freedom were fixed and could not be controlled by the participant. In the continuous non-immersive condition, the knob of the mouse was used to control the angular change of heading and pitch. In the discontinuous non-immersive condition, four mouse buttons were used to control displacements of the viewing orientation to the left, to the right, up and down.

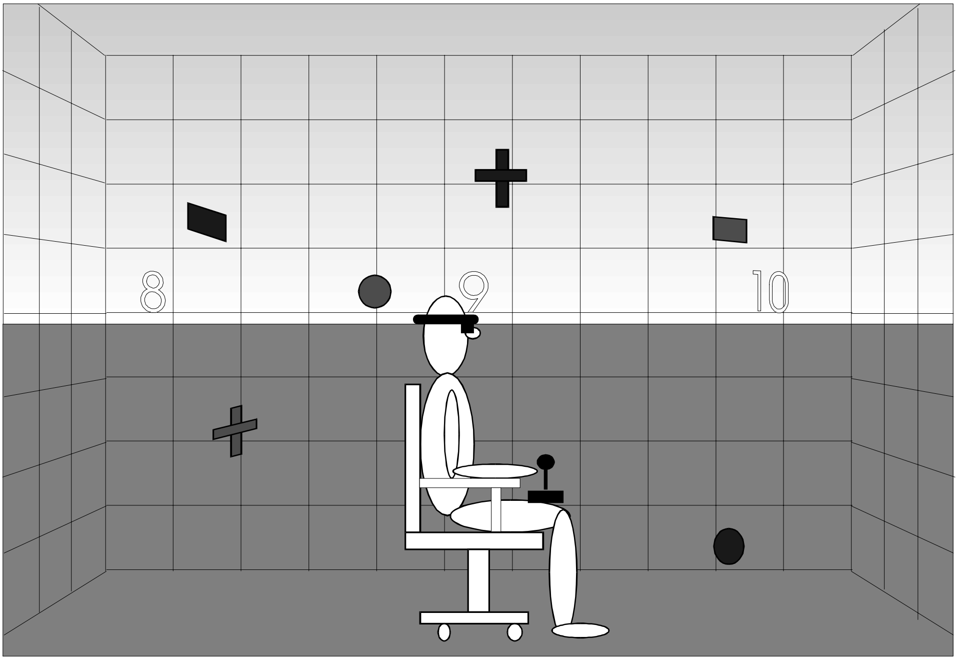


Figure 4.2: Impression of the Virtual Environment with a participant seated on a swivel chair equipped with an HMD and an input-device placed on the lap.

With continuous navigation, the images were updated to the current viewing direction of the observer by 20Hz, providing a fluent visual movement (Fig. 4.1). In the discontinuous condition, the area around the participant was segmented in 3 (pitch) x 12 (heading) = 36 sections. The size of each section corresponded to the geometric Field Of View (FOV) of the HMD so that no overlap existed between neighbouring sections. When the participant moved around immersively within a section, the same frozen image was seen, but as soon as the border between the two sections was crossed, the image switched instantaneously to the image corresponding to the next section.

Virtual worlds

Because a virtual world can only be learned for the first time once, four different but similar virtual worlds were created. The shape, the colour and the location of the six objects in each world were changed but the background remained the same. All four virtual worlds consisted of the following elements (Fig. 4.2):

- Six different objects of two different colours and three different shapes with a section of roughly 0.2 meters making them clearly visible in the scene (see Appendix B for a picture of the objects). The objects were located on an imaginary cylinder around the observer with a radius of 1.5 meters, extending between -33.7Ñand + 33.7Ñ in pitch from the viewpoint of the participant. Participants had to learn the location of these six objects during a learning phase and to search for them again during a test phase.

- A target object was displayed at a fixed position in the upper right part of the screen, indicating to the participant which object had to be searched for during a trial. This target corresponded as regards shape and colour to the object that had to be searched for, but it was slightly smaller.
- Four transparent walls each 5 meters long were placed perpendicular to each other around the participant, all with a grid of black stripes 0.3 x 0.3 meters apart. The walls provided optic flow while the participant was moving through empty sections with no objects.
- A horizon was made visible by the division between the green ground and the blue air.
- Numbers of 1 to 12 were placed just above the horizon outside of the transparent walls with grids to indicate the directions corresponding to the hours on a clock. This heading-taper was meant to aid orientation on the basis of recognition.

Apparatus

Participants were seated on a swivel-chair with a Head Mounted Display (HMD) on their head and equipped with a Polhemus head tracker. They also had a SpaceMouse™ fixed onto a small wooden board on their laps. Images were generated by an Onyx Reality Engine and were displayed using the Virtual I/O stereoscopic HMD. The geometric FOV was set to 30Ñ(H) x 22.5Ñ(V), resulting in a slight mimmification of the image. For more detail about the apparatus see Chapter 2.

Procedure

Participants were given written instructions about the nature of the experiment, the task and the procedure, after which they were allowed to ask questions. They were required to give written consent, and completed a questionnaire about their computer experience and their spatial orientation abilities.

Depending on the participant's speed of learning, the experiment lasted between 2½ and 4 hours during the morning or the afternoon. A break was given in between each session. Before starting each session, participants were allowed to familiarise themselves with the navigation interface for ten minutes in the same virtual environment without the six objects.

During the learning phase after each learning block, the HMD was removed and participants sat down at a table to answer a questionnaire that tested the quality of their CM. There were three types of questions illustrated by the next three examples: is the white square positioned above the horizon; is the white square positioned higher than the yellow cross; and is the white square closer to the red square than to the yellow cross? The questions were chosen as a gauge of the quality of the cognitive map because they provided no benefit for one condition as opposed to another.

Before starting to learn a layout, participants could look around in a virtual world for five minutes to get accustomed to controlling the navigation interface. The six objects were not visible during this period.

Scoring

Learning continued until a good CM was acquired. To test the quality of the CM, participants were given five written questions after each block, about the positions of objects in relation to each other. The blocks were repeated until a participant could answer a set of five new questions without error. *Learning duration* was defined by the number of learning blocks needed until participants could complete a full list of questions without error.

For the test phase, three time intervals were measured to indicate performance (Fig. 4.3):

- $T_{prepare}$: The time between the moment that a new target was shown and the moment that a participant pressed the start button to indicate that he remembered the object location. This time could be used by participants to retrieve a given target object from their CM and it could also be used in the fixed condition to determine the required route.
- T_{start}: Time (s) to start moving. This is the period between pressing the start button and the moment that the participant traversed the minimum angle of 30Ñ. This measure is especially important for indicating the time that it takes participants to determine their own location after random displacement in the test phase.
- N_{detour}: The number of trials in which participants turn at least 90Ñ more than the minimum required heading change to move to the target. This measure was taken as an indication of disorientation in the test phase.

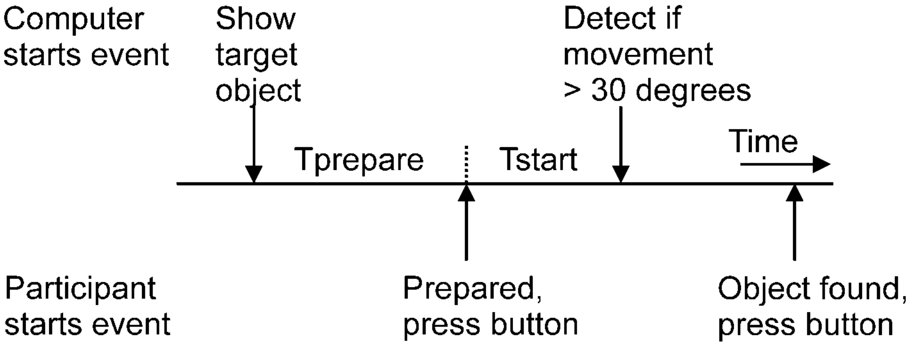


Figure 4.3: Procedure and dependent measures for a trial in the test phase

Apart from these objective performance measures, participants were given a questionnaire and were interviewed by the experimenter about the strategies that they used. At the end of the experiment participants were asked to rank the navigation interfaces from easiest (1) to hardest (4).

4.2.2 Results

Spatial knowledge acquisition

Our main research question was whether the acquisition of a CM was influenced by the type of interface or the type of displacement. During the experiment, it became clear that participants learned the VE faster in the later sessions. Therefore, the results from the learning phase were first analysed using a two-way ANOVA *session* (4 levels) x *navigation interface* (4 levels). The interaction and both the main effects were found to be significant (Table 4.1).

The mean learning duration (Fig. 4.4) shows a very strong effect of session, which was not anticipated prior to the experiment. A post-hoc Tukey test ($p<0.05$) shows that Session 1 differed from the later three sessions.

Separate analyses of Session 1 with a two-way ANOVA *interface* (immersive versus non-immersive) x *displacement* (continuous versus discontinuous) only shows a significant effect of *displacement* ($F_{1,12}=9.45$, $p<0.01$), indicating an advantage for continuous over discontinuous displacement (Fig. 4.4). However, when analysing the later sessions, no differences were found.

At the end of the experiment participants were asked to rank the navigation interfaces from easiest (1) to hardest (4). The immersive continuous condition (average rank 1.6) had significant advantages (Kruskal-Wallis ANOVA by Ranks: $H_{3,64}=12.62$, $p<0.006$) over the other conditions: immersive discontinuous (average rank 2.6), non-immersive continuous (average rank 2.8), non-immersive discontinuous (average rank 2.9). The pattern of rankings (Fig. 4.5) indicates that participants show agreement about the immersive continuous and non-immersive discontinuous conditions, but differ in their judgements on the other two conditions.

Table 4.1: Analyses of learning duration with two-way ANOVA session x navigation interface.

	df	F	p
session	3, 48	12.73	<0.0001
navigation interface	3, 48	2.89	<0.045
session x navigation interface	9, 48	3.09	<0.005

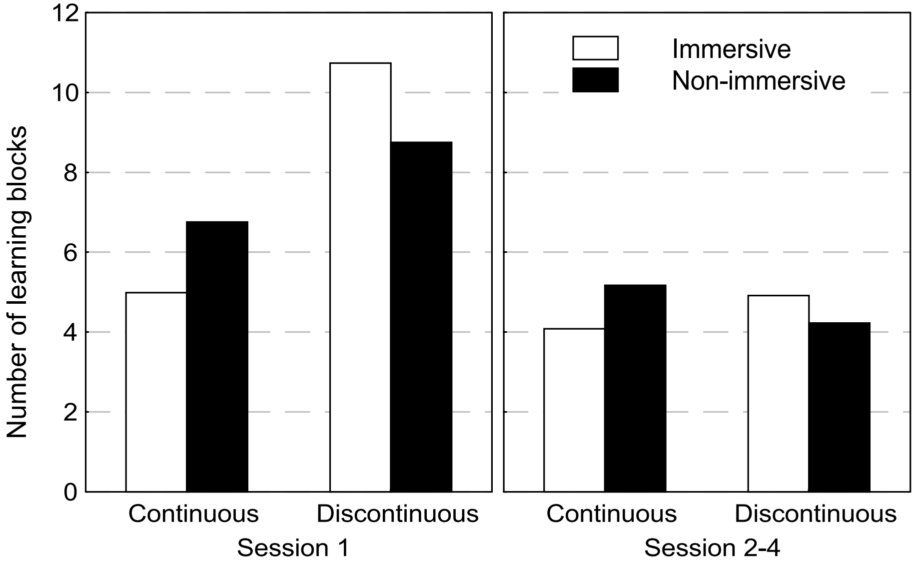


Figure 4.4: Mean duration of training for the four stimulus conditions for Session 1 (left) and Session 2, 3, and 4 combined (right). Each training block consists of six trials.

Testing

For the test phase the effects on $T_{prepare}$, T_{start}, and N_{detour} were analysed separately using three-way within-participants ANOVAs *interface* (immersive versus non-immersive) x *displacement* (continuous versus discontinuous) x *start-orientation* (fixed versus random). Furthermore, the effect of the factor *session* was tested separately with three one-way ANOVAs.

With $\mathbf{T_{prepare}}$, only the interaction between *interface* and *start-orientation* was found to be significant ($F_{1,15}$=5.3, p<0.04). However, the differences found in this interaction were very small. In the immersive condition, $T_{prepare}$ decreases from 3.1 seconds in the fixed treatment to 3.0 seconds in the random treatment. In the mouse condition, $T_{prepare}$ increases from 2.8 seconds in the fixed treatment to 3.0 seconds in the random treatment. A post-hoc Tukey test (p<0.05) only shows a significant difference between the fixed immersive and the fixed non-immersive condition.

More substantial is the decrease in $T_{prepare}$ over the four sessions (Fig. 4.6). $T_{prepare}$ decreases significantly from 3.6 seconds in the Session 1, to 2.9 seconds in the Session 2, to 2.8 seconds in the Session 3, and to 2.5 seconds in Session 4, with the standard deviation in all sessions being approximately 1.5 seconds ($F_{3,45}$=7.9, p<0.001).

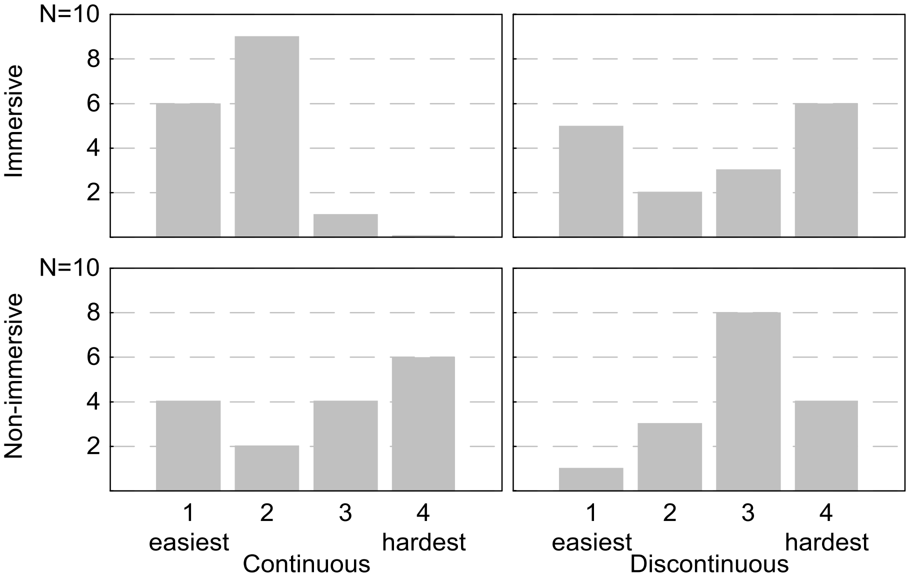


Figure 4.5: Number of participants (N) for each ranking of difficulty for the four navigation conditions.

$\mathbf{T_{start}}$ increased significantly from 1.3 seconds (s.d.=0.3 s.) in the fixed treatment to 1.9 seconds (s.d.=0.8 s.) in the random treatment (Fig. 4.7; $F_{1,15}=29.4$, $p<0.001$). No other significant interactions or main effects on T_{start} were found. No effect was found of the factor *session* ($F_{3,45}=0.3$, $p=0.85$).

With $\mathbf{N_{detour}}$, significant main effects were found of the factor *interface* ($F_{1,15}=7.7$; $p=0.014$) and of *start-orientation* ($F_{1,15}=8.7$, $p=0.01$). No other significant interactions or main effects were found. The effect of the factor interface shows that the percentage of trials with large detour is slightly higher in the immersive conditions (14%) than in the non-immersive conditions (10%). If put in the random orientation, participants became disoriented in 16% of the trials, which is twice as much as the 8% found for the fixed orientation (Fig. 4.7). No effect was found of the factor *session* ($F_{3,45}=0.4$, $p=0.74$).

Strategy

Some verbal reports made by participants suggest that object locations are stored hierarchically in the internal representation. Several participants grouped objects together mentally to aid recall. Many participants saw the virtual world as being built up of three layers on top of each other, especially in the discontinuous conditions. Certain participants reported knowing which layer the object was in, but not knowing precisely the orientation

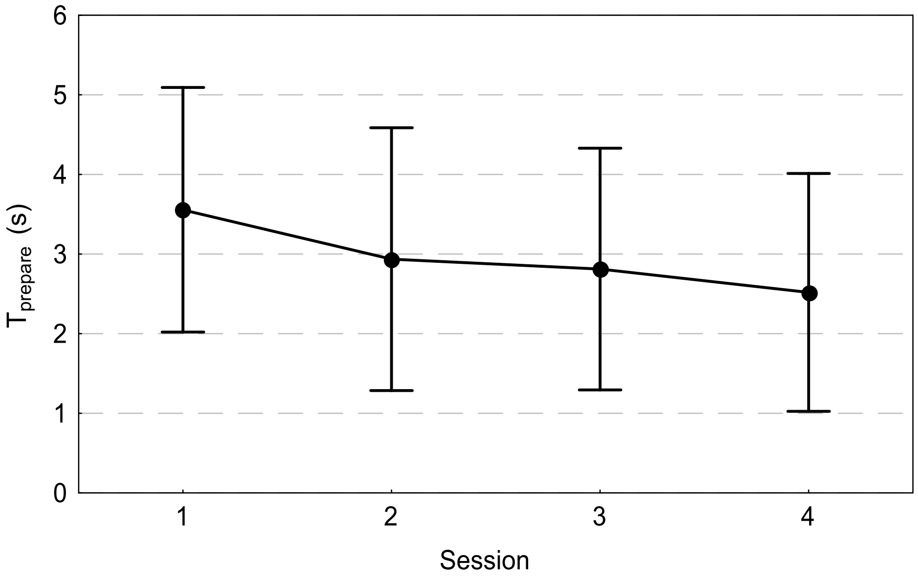


Figure 4.6: The effect of session on $T_{prepare}$ averaged over start-orientation, navigation interface, and participants. Whiskers indicate plus and minus one standard deviation.

in which the object was located. Interestingly, in the discontinuous conditions, participants reported knowing in which section an object was located, but did not remember the object location within the section.

Participants used very systematic search patterns, especially at the beginning of the learning phase in each session. As soon as more objects were known, participants started searching less systematically and moved more directly to the target objects that were known. Participants used three main search patterns, given below in order of their frequency of appearance:

- Layered search: participant turns around his axes, changing heading, while keeping his height or pitch constant for one full turn after which he/she moves to a different height and turns around again.
- Up-down search: participant moves up and down keeping heading constant after which he/she moves one FOV sideways and repeats the up and down movement. Although this search pattern involves more input effort to control the many direction changes, the participant in question maintained a good sense of direction because each time he passed the horizon he saw the heading-taper. A variation in the up-down search was the up-down zigzag movement: participants moved up and down while at

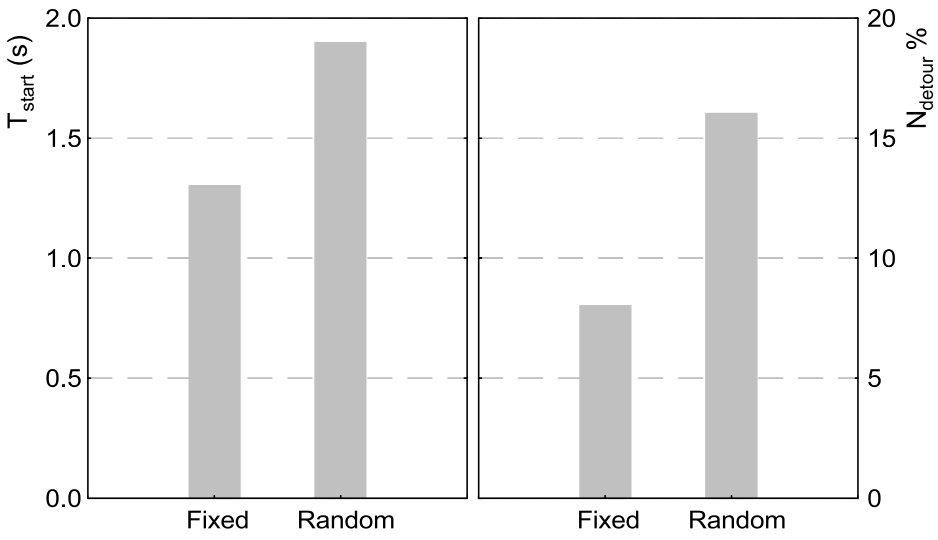


Figure 4.7: The effect of start-orientation (fixed or random) on T_{start} (left) and on N_{detour} (right).

the same time changing their heading in one direction. This zigzag pattern was only used by a few participants and only in the immersive continuous condition.

- Sector search: the participant stays in one quadrant looking around, changing both the heading and the pitch, after which the next quadrant is searched. Sector search was mainly used in the immersive conditions. Participants kept their bodies fixed while inspecting one sector with head movements, after which they moved to the next sector using their body.

Participants in the two non-immersive conditions primarily used the layered search pattern, mainly because it requires the least input effort. Participants just have to give the same input or a constant input most of the time to change heading, and they only have to change their input occasionally if they want to change to a different pitch angle. Participants in the two immersive conditions used slightly more up-down search than layered search, which was again done slightly more than sector search. No substantial differences were found between search patterns used in the continuous or the discontinuous displacement conditions nor were changes observed in the exploration strategy over the sessions.

4.2.3 Discussion

Spatial knowledge acquisition

The first phase of Experiment 6 was performed to test whether path integration contributes to spatial updating in cases where visual recognition can also be used. The navigation interfaces show clear differences in learning duration in Session 1. However, learning performance improves in later sessions from the point of view that all differences between conditions disappear in later sessions. This unexpected effect reduces the power of the experiment considerably, thus reducing the chances of finding significant results. In retrospect, the within-participants design proved to be unsuitable for this experiment. Three explanations can be formulated for the reduction of learning duration in the later sessions:

- Firstly, the cause for improvement may lie in a calibration of the path integration process. Experiment 5 in Chapter 3 has shown that path integration can easily be calibrated by putting participants in a virtual environment with visible recognisable objects. In the current experiment, participants might have used the heading-taper to calibrate their path integration process in Session 1. This could lead to an improvement of spatial knowledge acquisition in the later sessions.
- Secondly, similarities between the four different virtual worlds that were used may have facilitated visual recognition in the last three sessions. Although the six objects in each session were different, all four environments had identical backgrounds. Therefore, a part of the CM learned in Session 1 could be used by participants in later sessions to recognise their location. Recognising the background would reduce participants' dependency on the inaccurate path integration process.
- Thirdly, participants may have adapted to the learning task at the cognitive control level and have developed strategies to direct their attention in Session 1. Such task familiarisation would allow participants to learn the VE faster in Session 2, compensating for the degraded information that is provided by some of the spatial updating processes.

The last two explanations would also predict an improvement of the continuous immersive condition over sessions in this experiment, but this was not found to be the case. The lack of effect of session on this condition might be a ceiling effect. The literature offers little help when it comes to choosing between these hypotheses. A reduction of learning duration due to increased VE experience was something that was observed by Ruddle et al. (1997) in experiments where participants had to learn several virtual building layouts. However, they did not compare different navigation interfaces, nor did they offer an explanation for this effect. Chance et al. (1998) even found a completely opposite effect when they compared navigation interfaces that differ in the level of immersion. Only in the last of three similar sessions that participants had to perform, were significant differences found between the interfaces, but no effects were found in the first two sessions, or when all three sessions were combined.

Regarding only Session 1, where differences are apparent, only four participants are left in each group for comparison of the conditions. As expected, an increase in the duration of spatial knowledge acquisition was found with discontinuous displacement when compared to continuous displacement. A trend was found that suggests that immersive navigation might have some advantages over non-immersive navigation in the case of continuous displacement.

The subjective ranking of the four different stimulus conditions clearly shows the preference of participants for the continuous immersive condition and their dislike of the discontinuous non-immersive condition. However, care must be taken not to overrate these subjective rankings. Participants may not only rank the difficulty they have acquiring a CM when using the specific interface, but they may also rank other psychological factors like enjoyment or even physical effort while using the interface.

In the discontinuous immersive condition, participants reported experiencing difficulty because they could not tell precisely when the virtual viewpoint would be displaced to the next section if they moved their heads. Consequently, involuntary discontinuous displacement occurred, which may have disoriented the participants. In the non-immersive discontinuous conditions participants knew exactly when the discontinuous jump would happen because they pressed a button themselves. Nevertheless, participants who received the immersive discontinuous condition in one of the later sessions, did not show longer learning duration, although their familiarisation with this specific navigation interface was identical to that of their colleagues who received this treatment in the first session.

Exploration strategies

The results showed that most participants resort to systematic search patterns at first when exploring a new environment. Similar, Darken and Sibert (1996) have found that participants use search patterns that follow the visual guidance that is offered to them, like for instance a visual grid.

A clear advantage of a systematic search pattern is that the object is always found in a limited space of time. Furthermore it reduces the burden of remembering which locations have already been searched. Once part of the CM has been acquired, the systematic search patterns are abandoned, and participants start searching directly in the locations where they think the object might be located. The specific search patterns that participants choose depend mainly on the type of navigation interface, and not on the type of displacement. Participants seem to minimise the effort that is needed to control the input-device.

Testing spatial updating

In the second phase of Experiment 6, participants were placed in a fixed or random orientation in order to measure the efficiency of spatial updating after discontinuous displacement. Furthermore, the purpose of the tests was to see whether differences exist between navigation conditions in the first phase regarding the quality of the acquired CM. The time given to prepare for the displacement and the time needed for visual recognition of the own location after displacement were registered, as well as the percentage of incorrect route choices.

An increase in preparation time was expected for the fixed start orientation when compared to the random start orientation, because participants should spend time anticipating the displacement. The results show no evidence of such an effect. There was a small but significant interaction indicating that with a fixed start orientation, the time required to prepare is slightly larger for the immersive interface than for the non-immersive interface. The cause of this effect is unknown. The time needed to prepare decreases over the sessions, indicating that the retrieval of the object location from the CM and the determination of a route becomes more efficient with practice.

When looking at the other performance measures, clear differences were found between fixed and random start orientations. Placing participants in a random orientation temporarily disorients them. The recovery from this disorientation takes 0.6 seconds on average. Bowman et al. (1997) found an increase in search time for targets of 1.2 seconds after a discontinuous jump when compared to a continuous movement to an end position. Our participants may have been faster in reorienting because they were better trained or because the heading-taper in the VE facilitated visual recognition.

Apart from the difference in recovery time, the number of incorrect route choices made if someone is placed in a random orientation is much higher than if they are placed in a fixed orientation. This shows that jumping to an unanticipated location not only requires time to visually recognise one's location, but it can also lead to complete disorientation, at least when put under time pressure. No change in performance was found between sessions for the recovery time or for the number of correct route choices. This would suggest that visual recognition is a basic human skill, because no session to session improvement was seen.

A significant but small increase was found in the percentage of trials a large detour was made when going from non-immersive (10%) to immersive (14%) navigation interfaces. This difference is probably caused by an inherent difference between the immersive interface and the non-immersive interface in the contribution of measurement noise. With the immersive interface the viewpoint displacements are controlled by input from the head-tracker, which is noisy. The integration of this noise will add to the total detour. Whereas in the non-immersive conditions, participants can remain completely motionless in the VE, by simply not touching the input-device.

The results of the test phase provide no substantial evidence that differences may exist in the quality of the CM that is acquired with the different conditions. Only two small effects were found that involve the factors *navigation* or *displacement.* A small decrease was found in preparation time for the non-immersive condition, which could not be explained. Furthermore, a small increase in detour was found with the immersive interface, which can be attributed to measurement noise.

In Experiment 6 the disadvantage of discontinuous displacement was shown compared with continuous navigation. However, the expected difference between an immersive and a non-immersive interface in spatial knowledge acquisition was not shown.

In Experiment 7 evidence is again sought for this last effect. Furthermore, the experiment tests whether a calibration of the path integration process may be responsible for the disappearance of differences between conditions in the later sessions.

Three conditions are tested, all with continuous displacement: a non-immersive navigation interface, a non-immersive navigation interface with path-integration training beforehand, and an immersive navigation interface.

It is the calibration of path integration that is expected to be responsible for the improvement of performance in sessions found in Experiment 6. In that case the non-immersive interface with path-integration training beforehand should enable faster acquisition of spatial knowledge than the non-immersive group without training, at least in the first session.

Participants take part in two sessions, both with the same interface. If calibration of path integration during exploration of a VE was responsible for the improvement in learning duration in Experiment 6, there should only be a reduction in learning duration for the non-immersive condition because the other conditions already provide for a good quality of path integration.

To exclude the possibility of the background being learned in the Session 1 and influencing learning performance in a later session, completely new backgrounds will be used for consecutive sessions. If the immersive condition and the non-immersive condition with path integration still show improvement in learning performance from session to session, then the development of remembering strategies must be responsible for any improvement in session learning.

The test phase of Experiment 6 showed that being placed in a random orientation proves to be disorienting and it takes time to recover from this relying on visual recognition. However, no difference was found between a fixed start orientation and a random start orientation in the time needed to prepare. In Experiment 7 the tests are repeated, including a test in which the start orientation is not changed in-between trials. In that case, no time should be needed to prepare for the discontinuous displacement.

4.3 Experiment 7: The effect of path-integration training

4.3.1 Method

Participants

Thirty participants, eleven male and nineteen female, participated in the experiment. The participants, who came from different universities, ranged in age from 18 to 26 years, and had normal or corrected to normal eyesight. Participants gave written consent and were paid a fixed amount for their participation.

Task

Participants underwent a number of different tests and tasks: path integration test or training, two sessions of virtual world learning and testing, followed by another path integration test.

For path integration the same method and task were used as in Experiment 3 on path-integration training. Participants had to turn in a virtual forest as accurately as possible, with or without Knowledge of Results (KR) depending on the group they were assigned to.

The task and instructions during virtual world learning were identical to those of Experiment 6, although the training criterion, the virtual world and the procedure were somewhat different. Instead of using questions to test adequate spatial knowledge, the detour that participants make in reaching the target was used as a guage of their knowledge adequacy. Learning continued until eight out of the ten last performed trials showed a detour that was less than a quarter circle.

Design

A between-participants design was used with three navigation conditions: immersive, non-immersive, and non-immersive KR. Ten participants were assigned randomly to each condition, but care was taken to balance the number of males and females and the amount of experience with computer games as much as possible.

All participants had to perform 36 path integration trials identical to trials in Session 2 in Experiment 3. Only those in the non-immersive KR group received feedback about their performance (Section 3.2.1 for details). The session was chosen to be somewhat shorter than in Experiment 3 to reduce the duration of the experiment, while still giving adequate opportunity for path integration calibration for the non-immersive KR group. The path integration test was repeated at the end of the experiment for both the non-immersive groups to check for changes in performance, but without feedback for both the groups.

After the first path integration session, each participant had to learn the layout of two virtual worlds, of which the order of occurrence was balanced within each condition. After learning each world, two knowledge tests were performed, one test consisting of the same questions about object locations as in Experiment 6, and one test consisting of a drawing task. After these tests, the task of finding objects was tested from three different types of start orientations: a *fixed* orientation corresponding to the learning situation, a *random* orientation in which the heading was changed, and an *uninterrupted* start-orientation in which the start-orientation was equal to the end orientation in the previous trial so that no displacement occurred. The order of the three options was approximately balanced per condition. Each start-orientation was tested in a block of 30 trials where the first six were regarded as training and were not analysed. Participants were told which treatment they would receive.

Stimulus conditions

Navigation interfaces

Two different navigation interfaces were used that were identical to the continuous immersive and continuous non-immersive interfaces used in the previous experiment (Experiment 6). The non-immersive KR interface is identical to the non-immersive interface but this group was trained in path integration by providing performance feedback during path integration trials before learning the virtual worlds. The other two groups performed the same path integration trials but then without feedback.

The navigation interface for path integration trials is nearly identical but it is only the heading that is controlled. Participants in the immersive group of this experiment executed their trials while seated on a swivel chair, unlike in Experiment 1 and Experiment 2 where participants in the immersive condition were standing on their feet.

Virtual worlds

The visual stimulus during path integration trials, a virtual forest, was identical to that of Experiment 3. The virtual worlds that had to be learned differed from the worlds in Experiment 6, because similarities in both worlds needed to be reduced in order to avoid knowledge gained in one world being transferred to the other: The transparent walls with virtual grid and the heading-taper with numbers were not present. The heading-taper was removed to prevent non-spatial memory strategies involving arrhythmic ones being used to calculate the shortest route to a target. The transparent walls were removed because the square structure might reduce the need for path integration.

Instead, each world contained a large number of unique objects in the background, which could be used to recognise the current location. Both virtual worlds were constructed with enough objects so that there was always at least one unique object in view. The objects were for instance a water tower, a barstool, a cow, a bench, or a car. The six objects, the target object and the horizon are identical to Experiment 6.

Apparatus

A Kaiser ProView60 stereoscopic HMD was used with a physical FOV of 48Ñ (H) x 36Ñ (V), a binocular overlap of 100% with adjustable eyepieces. The geometric FOV was set to 30Ñ (H) x 22.5Ñ (V) in order to correspond to Experiment 6 and the eye disparity was set at 7.0 cm. To prevent participants from being distracted by light coming from underneath the HMD, the room in which the experiment took place was darkened.

Procedure

Participants were given written instruction on the nature of the experiment, the task and the procedure, after which they were allowed to ask questions. They were required to give written consent, and to fill in a questionnaire about their computer experience and their spatial orientation abilities.

The experiment lasted between two and four hours for each participant. Before starting all the experimental sessions, participants were trained in navigating using a tracking task. An irregularly moving star-shaped object had to be kept in view for at least one minute. In between sessions participants were given breaks.

Scoring

Learning duration was defined by the number of trials needed by participants to achieve the training criterion. After learning, a questionnaire with thirty questions regarding the object locations was administered. The questions were similar to those asked in Experiment 6 as a training criterion. The number of correct questions is a measure of the quality of the acquired CM.

For the other dependent measures see Experiment 6. A ranking of conditions was not possible because a between-participants design was employed.

4.3.2 Results

Path-integration training

As expected, analyses of absolute errors in path integration showed that errors for the non-immersive KR group (29°) are on the same level as those of the immersive group (26°), both being significantly lower than the errors of the non-immersive group (52°) ($F_{2,27}$=9.67, $p<0.001$). Comparison of path integration before and after VE learning for both the non-immersive groups show no differences ($F_{1,18}$=0.43, p=0.5) and no interaction between navigation groups and between testing before or after ($F_{1,18}$=2.11, p=0.11).

Learning

Contrary to what was expected, there was no significant effect of navigation interface on learning duration ($F_{2,27}$=1.76 $p<0.19$). The average learning durations for the three groups were 43, 49, and 32 trials for the non-immersive, the non-immersive KR and the immersive groups, respectively. Even when comparing the two non-immersive groups on the one hand with the immersive group on the other hand, no significant effect was found ($F_{1,28}$=3.06, $p<0.09$).

As was the case in the previous experiment, there was an effect of *session* indicating that participants learn the first layout more slowly, in 48 trials, than the second layout, in 35 trials ($F_{1,27}$=4.97, $p<0.03$). No interaction was found between *session* and *navigation interface*. We found no difference between the two different virtual worlds that were used ($F_{1,29}$=1.84, $p<0.19$). The two tests on paper that concerned the quality of the CM showed no significant differences between conditions. Participants answered only 22 out of 30 questions correctly, showing that their spatial knowledge after learning was far from perfect.

Testing

$T_{prepare}$, T_{start}, and N_{detour} were analysed with separate three-way ANOVAs on *session* (2 levels) x *navigation interface* (non-immersive, non-immersive-KR, immersive) x *start-orientation* (continuous, discontinuous-fixed, discontinuous-random).

With $\mathbf{T_{prepare}}$ only main the effects of *session* ($F_{1,27}$=4.53, $p<0.04$) and of *start-orientation* were found ($F_{2,54}$=4.74, $p<0.01$). As expected $T_{prepare}$ is lower for the Session 2 (2.5 seconds) than for Session 1 (2.8 seconds). For the different start-orientations $T_{prepare}$ is lowest for *random* (2.4 seconds), followed by *fixed* (2.7 seconds), and last *uninterrupted* (2.8 seconds). A post-hoc Tukey test ($p<0.05$) reveals that only the difference between the two most extreme values is significant.

For $\mathbf{T_{start}}$ a significant interaction between *session* and *start-orientation* was found ($F_{2,54}$=3.54, $p<0.04$) as well as a highly significant effect of *start-orientation* ($F_{2,54}$=28.57, $p<0.0001$). Only a marginally significant trend for *session* was found ($F_{1,27}$=3.73, $p<0.06$). Inspection of the interaction with a post-hoc Tukey test shows that T_{start} values are lower for Session 2 but only for the *random* start-orientation (a difference of 0.2 seconds). The main effect of start-orientation shows the lowest T_{start} for *fixed* (1.1 seconds), followed by *uninterrupted* (1.3 seconds), and last *random* (1.7 seconds). A post-hoc Tukey test ($p<0.05$) shows that all differences between start-orientations are significant.

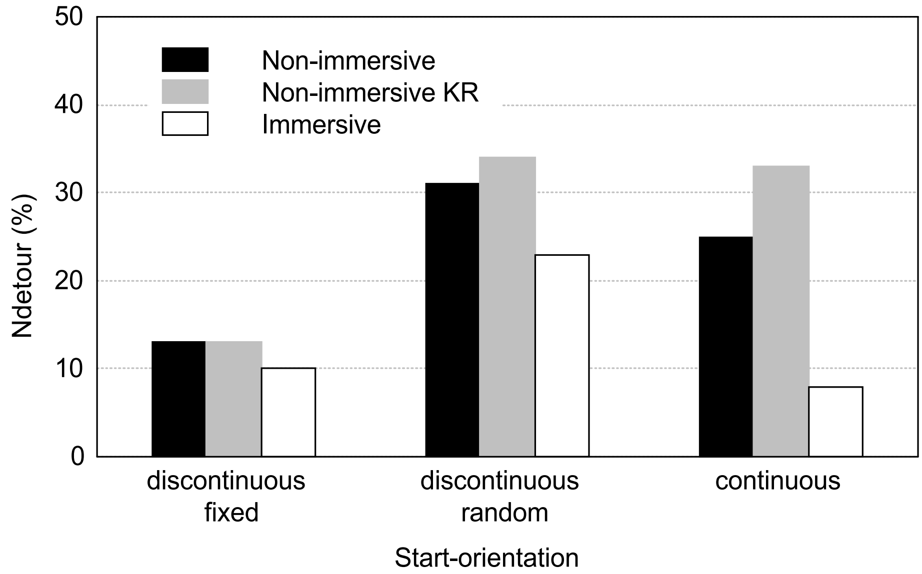


Figure 4.8: Average percentage of detour trials (Ndetour) for the interaction between the factors navigation interface and start-orientation, averaged over all participants.

For $\mathbf{N}_{\mathbf{detour}}$, a highly significant interaction between *navigation interface* and *start-orientation* was found ($F_{4,54}$=4.61, p=0.003) as well as a significant main effect of *navigation interface* ($F_{2,27}$=5.11, p=0.13) and a highly significant effect of *start-orientation* ($F_{2,54}$=29.98, p<0.0001). Inspection of the interaction with a post-hoc Tukey test showed a significant difference only in the *uninterrupted* start orientation between the immersive navigation-interface (10%) on the one side and the two non-immersive navigation interfaces (29% averaged over both the non-immersive interfaces) on the other side (Fig. 4.8). The percentage of detours is larger (34% at maximum) than the maximum percentage allowed in the learning phase (20%). The fact that participants did attain the training criterion is partly due to chance. Resampling statistics shows that with the training criterion that was used, participants with a failure rate of 34% need an average of 20 trials to reach the criterion. Our results showed learning durations that were well above this value (48 in Session 1, and 35 trials in Session 2).

4.3.3 Discussion

Experiment 7 was performed for three reasons. Firstly, a confirmation of hypothesis 2 was sought stating that there should be a difference between immersive and non-immersive navigation in the speed of learning a spatial layout. Secondly, it tested whether the improvement of learning from session to session that was found in Experiment 6 could be explained by a calibration of the path integration process. Thirdly, tests were performed to measure the efficiency of the spatial updating processes after different types of displacements.

Spatial knowledge acquisition

Contrary to what was expected, the results show no significant difference in learning duration between the immersive interface or the non-immersive interface, although there is a trend suggesting a possible advantage for the immersive interface. It is possible that the differences in spatial abilities between participants are so great that a possible effect of interface type is obscured.

Since there is not even a difference between the immersive and the non-immersive interface, it is not surprising that no difference was found either between the non-immersive interface with or without path integration training beforehand. The magnitude of this last effect should at least be smaller than the effect of immersion versus non-immersion thus making it harder to detect experimentally.

The result of the questionnaire done after learning showed that the acquired knowledge was incomplete. Only 73% of the questions were answered correctly whereas in Experiment 6 all the questions were answered correctly because the more stringent training criterion required this.

With all navigation conditions, improvement is found in learning performance during sessions. Calibration of path integration during exploration of the VE cannot be responsible for this improvement, because both the immersive condition and the non-immersive condition already allow accurate and calibrated path integration. Completely new virtual worlds were used for the two sessions including different backgrounds. Therefore, the spatial knowledge gained in Session 1 could not be used in Session 2. Learning performance improvement must be attributed to task familiarisation, allowing more optimal performance of cognitive control.

Testing spatial updating

The test phase provides some evidence that the quality of the CM that is acquired is higher in the immersive condition than in the two non-immersive conditions. High values for the percentage of detour were only expected for the random start-orientation. For the immersive group this is indeed the case. However, the two non-immersive groups also produce a high percentage of detour trials in the uninterrupted start orientation, contrary to what was expected. Apparently the participants in the non-immersive groups made errors when determining their own location or when determining the shortest route to the requested object in the uninterrupted condition. In both cases this points to a poor quality CM for the non-immersive groups.

Surprisingly though, the CM is adequate enough to enable good performance if a fixed start-orientation is used. Possibly the difference between performance with fixed and uninterrupted start-orientations was caused by a difference in the complexity of the choice of route. The fixed condition has only one possible start-orientation, whereas the uninterrupted condition has six different start orientation possibilities. This means that for the fixed condition there are only six different routes to choose from, and there is always only one single shortest route associated with a specific target object. For the uninterrupted condition there are thirty different possible routes to choose from, because for each of the six target objects five different start-orientations are possible. Apparently the process that determines the route is able to cope with the reduced quality of the CM in the easier fixed condition, whereas it is unable to reliably produce correct routes in the more difficult uninterrupted condition.

When looking at the time to prepare and the time to start moving for the different tests, the results confirm our expectations. The results show that the time to prepare is shortest for the random condition. In this condition, participants simply lack the information to determine a route beforehand. The preparation time does not depend on the type of interface.

As was also the case in Experiment 6, participants took longer to start moving if placed in a random start orientation than if placed in a fixed orientation. The extra time required to start moving is needed to visually recognise the own location and to determine the shortest route.

4.4 Conclusions

As yet, no sound confirmation has been found for Hypothesis 2. The difference in path integration performance between the immersive and the non-immersive interface does not lead to significant differences in learning duration. However, a clear trend in that direction is present in both experiments, and the subjective comparison of the conditions indicates that participants rank the immersive condition as the least difficult. Furthermore, some evidence was found that the quality of the CM is poorer after exploration with the non-immersive interface than with the immersive interface.

Navigating by discontinuous displacements disrupts the acquisition of a CM. Participants lack some of the motion feedback and need to use prior knowledge about the displacement to compensate for this loss.

If participants gain more experience with the exploration of VEs, then the time that is needed to acquire a CM decreases. This improvement in performance is most likely due to task familiarisation, which leads to more optimal performance of the cognitive control process.

Discontinuous displacement to a random orientation disorients participants temporarily, and it leads to an increase in the errors made in the choice of route after displacement. The set-up of the experiments in this chapter did not allow the role of cognitive anticipation to a discontinuous displacement to be fully isolated. The complexity of the choice of a route was not equal for the different displacement conditions, which led to a confounding variable. The role of cognitive anticipation in spatial updating will be explored more fully in the next chapter.

5 Navigation by hyperlinks[1]

Abstract

Hyperlinks in a Virtual Environment (VE) enable instantaneous displacement of the viewpoint over arbitrarily large distances. During hyperlinks path integration is not possible and spatial updating relies on visual recognition and cognitive anticipation. Two experiments investigate the efficiency of and the interactions between the spatial updating processes (Hypothesis 3). The first experiment investigates whether the type of hyperlink (varying according to the rotation, the direction, and the distance of the displacement) has an effect on visual recognition. The second experiment investigates the role of cognitive anticipation. Furthermore, both the experiments investigate the supposed advantage of an immersive versus a non-immersive navigation interface during the acquisition of spatial knowledge (Hypothesis 2).

Together, the results show an advantage of immersive navigation for the acquisition of a cognitive map. Recognition is not influenced by the type of hyperlink. Cognitive anticipation reduces disorientation after a hyperlink but not completely as with continuous navigation. The time needed for cognitive anticipation depends on the type of hyperlink. Some evidence was found that recognition of the current location is automatic.

5.1 Introduction

Displacement through a Virtual Environment (VE) can be made to go extremely fast by hyperlinks. In a hyperlink the viewpoint of the observer is displaced "instantaneously" over arbitrarily large distances from one graphics frame to the next. The hyperlink causes

---

[1] Parts of this chapter have been submitted as:

Bakker, N.H., Passenier, P.O, & Werkhoven, P.J. (submitted) The effects of immersion and discontinuous displacement on spatial orientation in virtual environments. *Human Factors.*

Bakker, N.H., Passenier, P.O, Werkhoven, P.J., Stassen, H.G., & Wieringa, P.A. (submitted) Navigation in virtual environments. IFAC conference 2001.

a discontinuity in our visually perceived position. The movement stimuli that are normally available for path integration are absent, which means that the path integration process registers no displacement. The location that is registered by path integration after a hyperlink still corresponds to the start-point of the hyperlink. The magnitude of the discrepancy of the path integration information with the actual location after a hyperlink is related to the type of hyperlink.

By definition, hyperlinks cause temporary spatial disorientation, which means that the internal representation of the current location no longer corresponds to the actual location in the VE. To recover from the disorientation, the internal representation of the current location must be updated to correspond to the new location. This updating depends on cognitive anticipation of the destination beforehand and on recognition of the location after taking the hyperlink.

Bowman et al. (1997) have shown that disorientation by a hyperlink takes time to recover from, at least if the exact destination of the hyperlink is not known beforehand. They measured the time participants took to locate targets after being translated (passively) either discretely or continuously to a location in the viewing direction of a known VE. Participants searched significantly longer for targets (approximately one second) after a discontinuous jump had been made than after making a continuous movement to their final position. Similar results were found in Chapter 4, where participants had to find a target after a random rotation.

In two experiments, the dynamics and the interaction of the processes involved in spatial updating during hyperlinks have been investigated. This may provide guidelines on the type of hyperlink that allows for good spatial orientation.

5.1.1 Hyperlinks

The spatial characteristics of a hyperlink displacement can be defined in the egocentric reference frame of the observer by the *hyperlink-distance,* the *hyperlink-direction*, and the *hyperlink-rotation* (Fig. 5.1). Hyperlink-direction is the translation direction relative to the start-orientation of the observer. Hyperlink-rotation is the difference between the end-orientation and the start-orientation of the observer.

Besides spatial relations between a start location and an end location of a hyperlink, numerous other relations could be defined. For instance, the start and end location may be close in appearance or in functionality. Closeness in appearance may affect recognition whereas closeness in functionality may affect cognitive anticipation of the destination. The current studies focus on the spatial characteristics of the hyperlink.

5.1.2 Visual recognition after a hyperlink

In the absence of cognitive anticipation, recovery from disorientation after a hyperlink depends solely on visual recognition of the new location. This recognition may involve the *loading* of a CM that corresponds to the new surroundings. Possibly, the CM that is retrieved from memory has to be *realigned* to correspond to the new orientation.

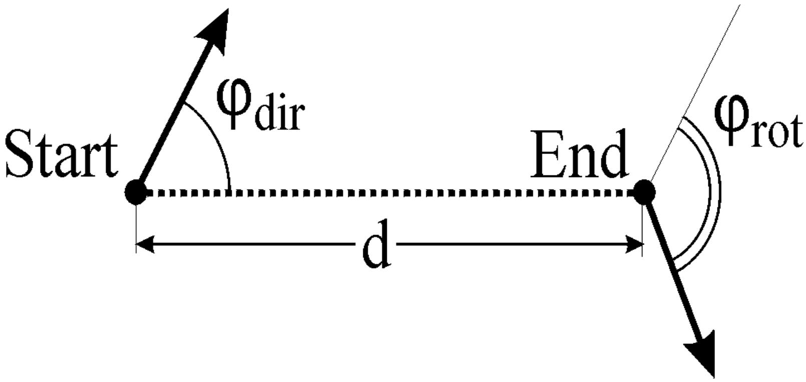


Figure 5.1: Relations between the start-point and the end-point of a hyperlink. The position and orientation both before and after a hyperlink are indicated with an arrow. The dotted line indicates the displacement by the hyperlink. Hyperlink-distance: “*d*”; Hyperlink-direction: “j dir”; and Hyperlink-rotation: “*jrot*”.

Although visual recognition does not depend on the information obtained from path integration, spatial updating may still suffer from interference created by conflicting location information from path integration. Farrell and Robertson (1998) and Farrell and Thomson (1998), showed that path integration during locomotion without vision is automatic and can only be ignored by deliberate cognitive processing. Similar results are found for path integration with a visual stimulus that does not allow recognition (May & Klatzky, 2000). Apparently, participants could not ignore information from path integration without performance decrement, despite being give explicit instructions to do so.

However, these results were all found in situations where the participant is moving and where recognition is not used at all. In our case the participant will be motionless during the hyperlink and recognition can be used to determine the current location upon arrival. Whether information from the path integration process still contributes to spatial updating at all in such a case is unclear. If path integration interferes with spatial updating depending on visual recognition after a hyperlink, the least interference would be expected if the path integration information was still partly correct after the hyperlink. In such cases, effects of hyperlink-rotation and hyperlink-distance are expected.

In the literature on mental displacements the observer does not move and recognition is not possible either. Participants are asked to point at an object as if standing at a different location. Looking at this literature, one sees that strong effects are found for the magnitude of rotation (Boer, 1991; Presson & Montello, 1994; and Rieser, 1989) and only weak effects are found for displacement distance (Easton & Sholl, 1995). Apparently a rotation is harder to cope with than a translation. If interference is caused by path integration in the case of a hyperlink, the strongest effects will probably be found for hyperlink-rotation. For hyperlinks with no hyperlink-rotation at least the orientation information provided by path integration is correct and less interference is expected.

Besides this interference effect yet another effect can be suspected that might influence spatial updating through visual recognition. The efficiency of access to information that is stored in the CM depends on the viewpoint of the observer in the environment. Remembering the location of objects depends on the current location of the observer (Boer, 1991; Easton & Sholl, 1995; and Sholl, 1987). This effect of relative object-bearing is that objects that are located more to the front of the observer are remembered faster than objects located more to the side, and even more than objects that are located behind the observer. This means that when trying hard to remember where an object is located it is actually beneficial to physically turn around in order to face possible directions.

Recognizing a new location after taking a hyperlink also involves using the CM to find the location corresponding to the recognised view in memory. The direction of the hyperlink destination in relation to the original viewing direction may well play a role in the recognition of the new location. If this is the case, an effect of hyperlink-direction is expected. A hyperlink-direction with a small angle should lead to smaller recognition times after the hyperlink.

Seemingly contradictory to these expectations is the fact that in the literature on mental displacements no effects of displacement direction are found (Easton & Sholl, 1995; and Rieser, 1989). However, the experiments that are reported were not set up to measure displacement direction. In these experiments, displacement direction is always confounded with relative object direction. Participants were in the middle of an array of objects and had to imagine a translation to one of the objects and point at another object. A mental displacement to the back (presumably more difficult than displacement to the front) is always combined more with pointing at objects located to the front (which is easier than pointing to the back), and vice versa. The two opposite effects of relative object-bearing and of mental displacement direction might have cancelled each other out.

The first experiment in this chapter investigates the influence of the relation between the start and the destination of a hyperlink on recovery from disorientation by recognition. Furthermore, the possible advantage of an immersive over a non-immersive navigation interface during the building of an internal representation is again investigated.

Participants first had to learn the layout of a VE and the location of objects placed in it using either an immersive or a non-immersive navigation interface. Learning duration is expected to be shorter if immersive navigation is used. All navigation in this phase was continuous and did not involve using hyperlinks.

After learning, participants are tested in the same layout with various hyperlinks. The time needed to update the internal representation of the current location after a hyperlink is measured by registering response latency in pointing at the previously learned objects. The hyperlinks that are tested vary in hyperlink-direction and in hyperlink-rotation. To increase the difficulty of cognitive anticipation different hyperlink-distances are also used, and all different hyperlinks are tested in a randomly mixed order.

Path integration is expected to interfere with the spatial updating by visual recognition. With small hyperlink-rotations the discrepancy between path integration information at the start and at the destination is smaller than with larger hyperlink-rotations. Therefore, smaller pointing latencies are expected in this case.

Furthermore, the relative direction of the hyperlink destination is expected to influence the speed with which the destination is recognised. Larger hyperlink-directions are expected to increase the pointing latency. The data are also used to verify the effect of relative object-bearing, on which this hypothesis was based. Objects located to the front should be remembered faster than objects located to the back (Boer, 1991; Easton & Sholl, 1995; and Sholl, 1987).

5.1.3 Cognitive anticipation to a hyperlink

When using hyperlinks in an interface, the destination of the displacement will often be known beforehand. Cognitive anticipation can then be used to prepare for the displacement, possibly reducing the effects of disorientation after the displacement. Indeed, the results presented in Chapter 4 showed that if the destination of displacement at the start of a trial was known, then participants would start to move faster to their target's location, when compared to starting to move after a random displacement. However, cognitive anticipation might not be able to completely compensate for the lack of path integration information. Bowman et al. (1997) compared a hyperlink in the viewing direction with continuous movement. Although participants could at least partially predict their destination, there was still considerable disorientation when compared to continuous movement. The literature offers no insight into whether disorientation by a hyperlink can be completely reduced when given complete knowledge beforehand of the destination. Taking a hyperlink with complete knowledge of the destination beforehand may still be different from arriving at a place after continuous navigation, which allows for recognition and path integration including perceptual anticipation.

The process of cognitive anticipation to a location after a hyperlink may be similar to the processes involved in mental displacement. For both the tasks an imaginary destination is given. If the cognitive anticipation process operates similarly to the processes involved in mental displacement, one would expect the same effects for the duration of cognitive anticipation to hyperlinks with different characteristics as for the duration of cognitive anticipation. This would mean that a strong effect of hyperlink-rotation should be present in accordance with mental rotation (Boer, 1991; Presson & Montello, 1994; and Rieser, 1989), and a weak effect of hyperlink-distance should be present in accordance with mental translation (Easton & Sholl, 1995).

Research on mental displacement shows that it is quite hard for participants to imagine a displacement that is not accompanied by actual movement of the body (Boer, 1991; Easton & Sholl 1995; May, 1996; Presson & Montello, 1994; Rieser, 1989; Wang & Simons, 1999; and Wraga, Creem & Proffitt, 2000;). Therefore, cognitive anticipation may also place high demands on limited resources, thereby inducing a high mental workload that may cause interference with other tasks. If using a VE as an operator interface, the task of an operator does not only involve the retrieval of a single item of information. Information from multiple sources needs to be gathered and integrated for use in decision making. This means that other data has to be remembered while searching for more information. If cognitive anticipation places a high demand both on attention recourses and memory recourses it may interfere with this information integration.

After displacement by a hyperlink, visual recognition is also possible. In principle, recognition is no longer needed because cognitive anticipation was already used to establish the current location after the displacement. However, the logical superfluousness of recognition does not mean that the process is not executed, and recognition may interact with cognitive anticipation. During natural locomotion, we use recognition repeatedly to compensate for the increasing errors in path integration. We are certainly not aware of this updating and it seems only fair to assume that little mental effort is needed to perform this visual recognition.

The information established by cognitive anticipation might be poor quality or hard to use. Recognition might still prove beneficial and strengthen the information about the current location. The cognitive anticipation beforehand may facilitate the recognition process making it almost instantaneous.

If recognition is triggered automatically, even without any need for the process, it might even slow down spatial updating. To see whether recognition is triggered automatically a conflict can be introduced experimentally between the information provided by the cognitive anticipation process and the recognition process. A viewpoint can be provided after displacement that does not correspond to the viewpoint that was anticipated. If participants are able to control the execution of recognition they should be able to ignore the inconsistent viewpoint and there should be no performance decrement in such cases.

Although recognition might be triggered automatically, cognitive anticipation is not an automated process, but one that requires deliberate intention to execute. However, once the process is activated it might be hard to stop. This last hypothesis is tested by asking participants to use cognitive anticipation while not giving them enough time to complete this process before the hyperlink displaces them. If the process cannot be stopped and if it requires more time, then spatial updating could be slowed down because participants are still busy anticipating while they are already displaced.

To summarise, four questions were raised pertaining to spatial updating and hyperlinks.

1. To what degree is cognitive anticipation able to compensate for the disorientation caused by hyperlinks?
2. Does cognitive anticipation follow the same dynamics as mental displacement?
3. Is recognition triggered automatically and how does it interact with cognitive anticipation?
4. What is the mental load involved in the spatial updating processes?

To answer these questions, a second experiment was executed consisting of two parts. In the first part, the same layout is learned as in the first experiment of this chapter, using either an immersive or a non-immersive navigation interface. Again, the possible advantage of an immersive over a non-immersive navigation interface during the building of an internal representation is investigated. In the second part, participants have to point at objects in the previously learned layout after different types of continuous and discontinuous displacement to one of the rooms has taken place. The different displacements allow different combinations the spatial updating processes, these being the path integration process, the recognition process, and the cognitive anticipation process. By administering the pointing task right after displacement the effectiveness of the spatial updating processes can be measured.

The response latency for pointing is taken as a measure of the time that is needed to update the internal representation of the current location. The possibility for cognitive anticipation is provided by showing participants numbers that tell the exact destination of the hyperlink. After cognitive anticipation, participants initiate the discontinuous displacement themselves, allowing the time needed for cognitive anticipation to be measured. To isolate cognitive anticipation from recognition, participants are also tested without recognition being possible by using a black screen in front of the participant after displacement. A double-task is introduced requiring participants to remember three items and their location of presentation on the screen.

Given earlier findings, cognitive anticipation should at least partly compensate for the disorientation by a hyperlink. If cognitive anticipation alone is sufficient to determine the current location, there should be no decrement in performance when compared to continuous navigation. The recognition process is expected to be automatic, which should decrement performance if an inconsistent viewpoint is shown. Cognitive anticipation is expected to show an increase in the workload experienced, which is not expected for the other spatial updating processes.

The hyperlink displacements vary in hyperlink-rotation and hyperlink-distance. It is expected that hyperlink-rotation and hyperlink-distance will have an effect on the duration of cognitive anticipation, thereby showing similar dynamics to imaginary displacement.

5.2 Experiment 8: Hyperlinks without cognitive anticipation

5.2.1 Method

Participants

Sixteen participants were used in the experiment, eight males and eight females. The participants, who came from different schools and universities, ranged in age from 17 to 28 years, and had normal or corrected to normal eyesight. Participants gave written consent and were paid a fixed amount for their participation.

Task

The experiment consisted of different phases in which either a part of the layout of the VE had to be learned or disorientation was tested following a hyperlink displacement in the learned layout. While learning, participants were instructed to learn the layout of objects in the VE **as fast as possible**. Participants could move freely in the VE using the navigation interface that was assigned to them. At fixed time-intervals during learning, participants' knowledge was tested by letting them point at objects **as accurately as possible**. Pointing at objects was chosen as a task because this requires participants to have knowledge both of the environment and of their current location. Once the VE was learned, the disorientation caused by hyperlinks was tested by letting participants point at objects **as fast as possible**.

Design and analysis

The factor *navigation interface* (immersive versus non-immersive) was varied between participants. Participants were assigned randomly to either the immersive navigation group or the non-immersive navigation group with the restriction that an equal number of males and females were assigned to each group. Both groups performed the same training and three tests: *spatial knowledge acquisition test,* a *baseline disorientation test*, and a *hyperlink test*:

1. In the *spatial knowledge acquisition test* the time needed by participants to learn the layout of objects in the VE depending on the factor navigation interface was measured. The test continued until full knowledge was gained of the spatial layout. Analysis was performed with a one-way ANOVA on the between-participants factor *navigation interface* (2 levels). Furthermore, the factors *computer experience* and *gender* were analysed with separate one-way ANOVAs.
2. The *baseline disorientation test* was carried out to verify whether an increase in disorientation after a hyperlink can be measured. Hyperlink-rotation was varied within participants. The pointing latency after a zero hyperlink-rotation was compared with pointing latencies after hyperlink-rotations of 90 and 180 degrees. Participants received a total of 45 trials in this test, consisting of the combination of 3 (start-orientation) x 3 (end-orientation) x 5 (objects). Analysis was performed with a two-way ANOVA with the between participants factor navigation interface (2 levels) and the within participants factor hyperlink-rotation (2 levels).
3. In the *hyperlink test* the independent variables are *hyperlink-rotation* (three levels: 0, 90, and 180 degrees), *relative object-bearing* (three levels: front, side and back) and *object* (5 levels for each room). Participants received a total of 243 pointing trials in this test, consisting of the combination of 3 (start-orientation) x 9 (end-orientation) x (9) (5 objects in same room + 4 objects in different room).
    - *Hyperlink-direction and hyperlink-rotation* was manipulated to see whether interference of path integration could be found in spatial updating after a hyperlink. Hyperlink-rotation was only compared for trials with equal end-orientation. Hyperlinks with end-orientation opposite an entrance were excluded, because they could not be reached by a jump with a hyperlink-rotation of 180 degrees, since the direction facing the entrance was not used as a start-orientation. Analyses was performed with a two-way ANOVA with the between participants factor navigation interface (2 levels) and the within participants factor hyperlink-rotation (3 levels).

- *Relative object-bearing* was manipulated to verify an effect found by Sholl (1987) and Easton and Sholl (1995). *Relative object-bearing* is the angle between the line connecting the observer with an object and the line straight-ahead in the direction faced by the observer. Data for relative object-bearings are collapsed for left and right directions and are divided into three categories containing an equal number of trials: front (< 60Ñ), side (between 60Ñ and 120Ñ degrees), and back (between 120Ñ and 180Ñ). Analyses was performed with a two-way ANOVA with the between-participants factor navigation interface (2 levels) and the within-participants factor *relative object-bearing* (3 levels).
- *Object* is tested to get an indication of whether differences occur in the recollection of specific objects. Analyses are performed with three separate ANOVAs, one for each room in the VE with the between participants factor *navigation interface* (2 levels) and the within participants factor *object* (5 levels).

In the case of each analysis, the data were collapsed over all factors not included in the specific analysis. The means were determined for each participant for the different factor-levels.

Virtual world and stimulus conditions

Virtual world

The virtual world consisted of four square rooms measuring five by five meters, connected on one side with a corridor (for layout see Fig. 5.2, for typical views see Fig. 5.3). All walls and floors were textured. On each of the walls of the room, except the ones with an entrance, a unique poster was hanging. Each poster contained a photograph of a different facility used for experiments at our institute. Five objects were placed in each of the four rooms, and they were evenly distributed in a fictitious circle around the centre of the room. All twenty objects were unique and distinctive, and were modelled on well-known household or office appliances (Appendix B). The objects were scaled to fit approximately in a 30 cm cube, and were placed on identical pedestals 1.5 meters high. The assigning of objects to a location and a room was done randomly to ensure that no grouping of similar objects would occur.

Object pointer

During testing participants used a graphic object pointer to indicate the bearing of the objects in relation to themselves (Fig. 5.3). Participants used the input-device to displace the dial of the object pointer to indicate the bearing of the object in relation to themselves. The mapping of object-bearing on the horizontal plane to dial direction in the vertical plane was identical to normal computer-mouse mapping. For the sake of clarity, pointing at an object behind one was established by pointing the dial downwards. The displacement of the dial into any direction was instantaneous. Because the dial did not move gradually, no extra time was consumed for larger dial changes, which would otherwise have affected the measured response latency. During pointing all the objects were removed from the scene and participants were no longer able to move in the VE.

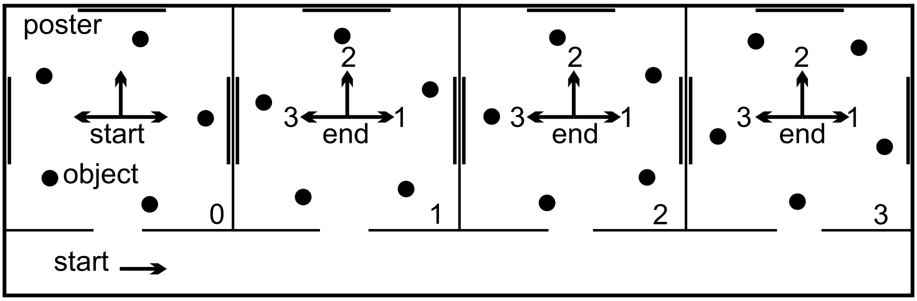


Figure 5.2: Layout of the virtual environment. Rooms are indicated with numbers 0, 1, 2, and 3 given in the lower right corner of each room. Each room contains five unique objects and a unique poster on each wall, except for the wall leading to the corridor. Arrows indicate the positions and facing directions for the start-points and the end-points of the hyperlinks. The three start-points in Room 0 are used in Experiment 8, and the start-point in the hallway is used in Experiment 9.

Navigation interface

To explore the virtual world, participants either used an immersive navigation interface or a non-immersive navigation interface. With both interfaces, participants were seated on an ordinary office swivel chair, equipped with an HMD to view the VE. An input-device was attached to a wooden board and placed on the participant's lap. The input-device that was used to control movement and to provide pointing responses, was aligned with the chair.

In the immersive navigation interface, the three viewpoint rotations were slaved to the head-motions of the participant by using an electromagnetic head-tracker attached to the HMD. Viewpoint translations in the horizontal plane were controlled with the input device. The direction of translation was relative to the orientation of the chair, which was determined using a second position-tracker. The chair therefore served as a virtual vehicle and participants could look around freely while moving independently of the translation direction of the virtual vehicle.

In the non-immersive navigation interface, participants controlled both the viewpoint rotations and the viewpoint translations in the horizontal plane with the input-device. For both the interfaces, the direction of the rotation and translation of the input device corresponded intuitively to the direction of translation and rotation of the viewpoint in the VE. The knob of the input-device has to be handled by the participant as if he is holding his own head, pushing and turning the knob to establish the desired change in viewpoint.

Apparatus

Images were generated by an Onyx Reality-Engine with Multi-Channel-Option manufactured by Silicon Graphics. The refresh rate was 60 Hz. and the update rate varied between 5 and 20Hz depending on the viewpoint in the VE.

A Kaiser ProView60 stereoscopic HMD was used with a physical FOV of 48Ñ (H) x 36Ñ (V), and a binocular overlap of 100% with adjustable eyepieces. The geometric FOV was set to 60Ñx 45Ñ (horizontal x vertical) resulting in a modest mimmification of the image. Head orientations were tracked using Polhemus Fastrak™. The input-device that was used was the SpaceMouse™ (Virtual Technologies). For more details about the apparatus used see Chapter 2.

Procedure

Before the experiment started a questionnaire had to be completed concerning general orientation ability and computer experience. The experiment consisted of four phases: training of input-device and object pointer; baseline disorientation test; virtual world learning; and final hyperlink test. The total duration of the experiment was approximately four hours and participants could rest in between the different phases.

Phase one served three purposes: to train the control of movement; to train the use of the object pointer; and to learn the location of objects in Room 0 in order to prepare for a baseline test in the next phase. The control of movement with the input-device was trained in a virtual hallway with left and right turns of 90 and 180 degrees (Fig. 5.3). To complete this training, participants had to walk along the entire hallway without bumping into walls. Next, the use of the object pointer was trained in Room 0 while at the same time the locations of the five objects in that room were learned. Participants could alternately inspect the object positions by looking around freely and point at a requested object using the object pointer. The participants' location was fixed to the centre of the room and movement was limited to changes in orientations. Participants themselves indicated when they wanted to point at an object. They did this by pressing a button on the input-device. After pressing the button, a large yellow arrow was shown in front of one of the three walls with a poster, to indicate the start-orientation. As soon as the participant had moved in the required direction, the object pointer was shown with the requested object in front of it. Pointing was repeated until participants could indicate the direction of all surrounding objects from three different orientations within a margin of plus or minus 45 degrees.

In **phase two**, the *baseline disorientation test* was held to verify whether there is indeed an increase in response time after displacement by a hyperlink in comparison with the response time after no displacement. In the test, participants' latencies were measured for pointing to the five objects in Room 0 that had already been learned in the previous phase. Just before pointing, participants were either displaced to one of the other two possible orientations in Room 0 or they were not displaced at all. The different trials were administered randomly, so participants could not anticipate whether they would be displaced or what the end-orientation would be in the case of displacement. Between pointing, participants had to move themselves to one of the three start-orientations, except in cases when the start-orientation of a new trial corresponded to the end-orientation of a previous trial. After reaching the start-orientation, participants were moved

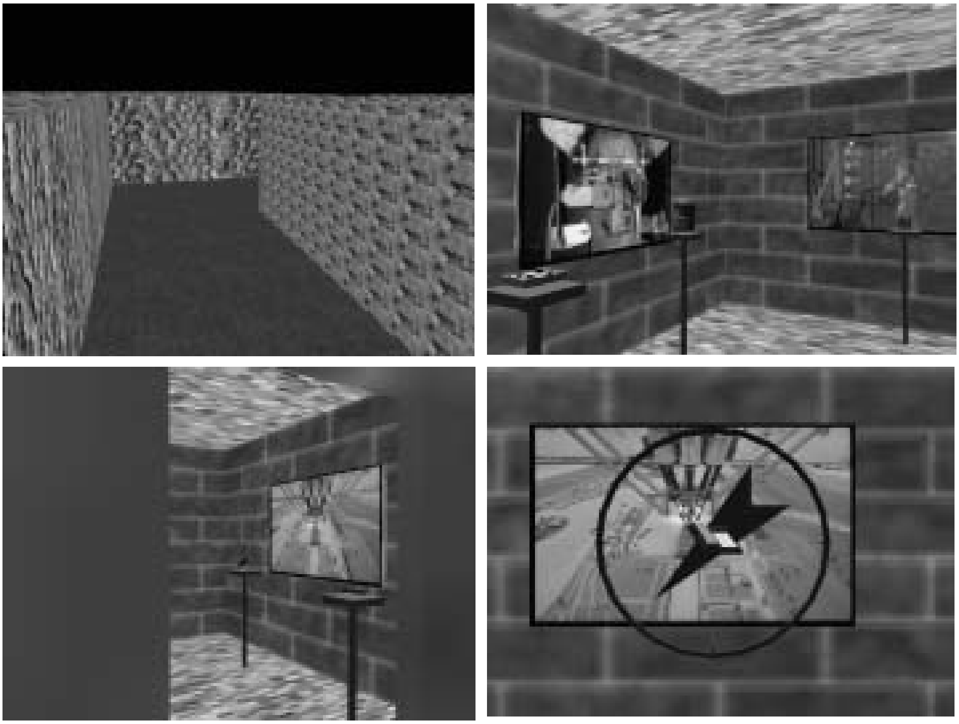


Figure 5.3: Typical views: top-left: hallway for the training of the navigation interfaces; top-right: inside Room 1 during spatial knowledge acquisition; bottom-left: Room 2 seen from the hallway looking through the entrance during spatial knowledge acquisition; bottom-right: pointing: the dial with in front of it the telephone with a poster in the background.

discontinuously or not (one out of three cases) and the pointer and requested object were shown. Participants were instructed to point as fast as possible at the target object. During the entire baseline test the objects were removed from the room.

In **phase three**, the duration of *spatial knowledge acquisition* was tested to see whether this is influenced by immersive versus non-immersive navigation. Participants were allowed to forget the objects in Room 0 and had to learn the locations of the remaining fifteen objects located in Rooms 1, 2, and 3 using either the immersive navigation interface or the non-immersive navigation interface to explore the VE. To be able to register the duration of learning, the exploration was interrupted at fixed time intervals to test whether participants had complete knowledge of the virtual world. In the test, participants had to point eighteen times at random target objects from random locations, all within a margin of 45 degrees or else learning resumed with free exploration. After each test the participant was informed of the number of errors made. The number of learning blocks required was used as a gauge to compare speed of learning in both the navigation conditions. The first training block started with an extra five minutes of free exploration in which the objects were not yet visible so participants could focus on learning the global layout of the rooms and the posters.

In **phase four**, the *hyperlink test* was administered to see whether there is evidence of interference in spatial updating caused by path integration. In each trial, participants started in the start hallway outside Room 0 and had to move themselves to one of the three start-points in Room 0, taking as much time as they needed. A disk in the centre of the room indicated the position of the start-point, and an arrow indicated the orientation of the start-point. A snapping function assured that participants could no longer move once they reached the start-point and were facing the required direction. Two seconds after the start-point was reached, the participants were displaced by the hyperlink to one of the nine end-points in one of the three other rooms. As soon as the end-point was reached, the pointer and the target object appeared. Participants then had to indicate the direction of the target object relative to their viewing direction as fast as possible. After target indication, participants were moved with a hyperlink back to the hallway to start the next trial.

During all phases of the experiment in which participants had to point at a target object, the twenty objects that were learned were removed from the scene.

Scoring

Before the experiment started a small questionnaire was done, in which participants' experience with computer games was assessed. During the VE learning phase, the dependent measure is the number of blocks of 90 second exploration, needed before participants reach the criterion of pointing eighteen trials all with errors of less than 45 degrees.

For the baseline test and the hyperlink test, the dependent measure is latency for pointing at objects, or the time between first presentation of a target object (directly after the hyperlink) and the participant's response by setting the dial of the object pointer and pressing a button. This measure can only be used to compare factors, and not in an absolute sense because the pointing latency also includes the time required to remember where the object is located and the time needed to use the object pointer.

Before collapsing and analysis, the data from each participant were cleaned. Only data regarding pointing at objects in the same room are analysed because the accuracy in pointing at objects depends on the distance between the participant and the object. All data points in which the participant did not move the pointer for object-bearing were removed. After that, the outliers for each participant were removed (values larger than four times the standard deviation around the participant's mean).

With our equipment the rendering performance depends on the viewpoint and viewing direction in the VE. Because rendering delays affect the pointing latencies, all tests are performed on datasets containing the same hyperlink end-points.

5.2.2 Results

Seven participants were unable to complete the experiment and were replaced to maintain a total of sixteen participants. Three of these participants fell ill, all during navigation with the immersive interface. Three participants were not able to learn the layout of the objects with the non-immersive interface within the limit of twelve training blocks. One participant was not able to adequately control movement through the virtual hallway that was used for training.

Spatial knowledge acquisition

The results of the virtual world learning phase show only a marginally significant effect of navigation condition ($F_{1,14}$=3.54, p<0.08), with a trend showing an advantage for the immersive group (5.6 learning blocks, with a std.dev. between participants of 1.85) over the non-immersive group (7.8 learning blocks, with a std.dev. between participants of 2.60). Participants who play computer games frequently (>1 time per month, N=11) learn the virtual world faster than participants who do not frequently play computer games (4.8 versus 7.6 learning blocks, respectively; $F_{1,14}$=5.9, p<0.03). No differences were found between men and women ($F_{1,14}$=0.3, p<0.6). In retrospect, the frequent and infrequent

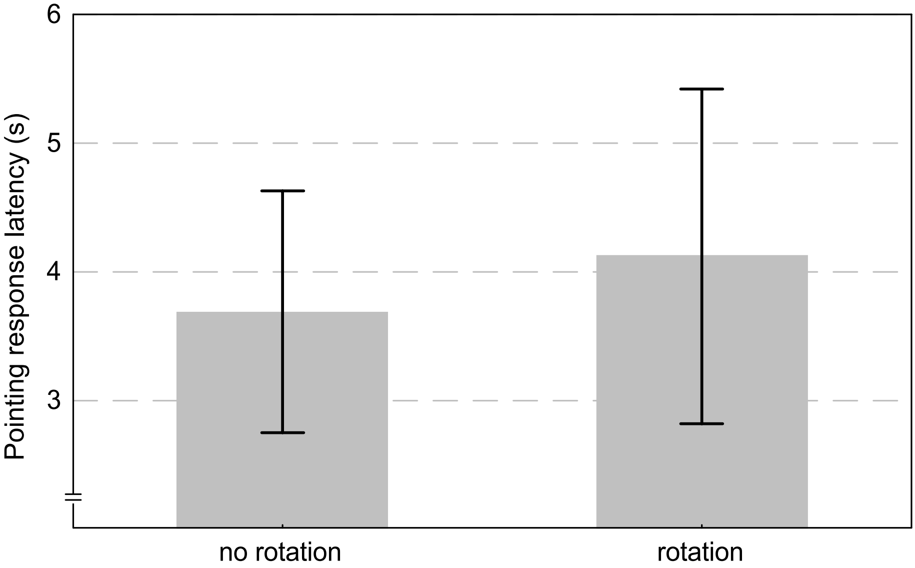


Figure 5.4: Mean pointing response latencies averaged over participants and navigation group, after no-rotation or after a rotation in the baseline test. Whiskers indicate plus and minus one standard deviation, which corresponds to the variation between participants.

game players were almost equally divided between the immersive group (6 frequent and 2 infrequent) and the non-immersive group (5 frequent and 3 infrequent).

Baseline test

Pointing after a discontinuous rotation, 4.1 seconds, is slower than pointing after no rotation, 3.7 seconds (Fig. 5.4; $F_{1,14}=6.6$, $p<0.02$). No significant main effect or interaction was found of the between-participants factor *navigation condition*.

Hyperlink tests

Hyperlink-direction

No effect was found for *hyperlink-direction* ($F_{2,28}<1$). The means for the different hyperlink-directions of 0, 90, and 180 degrees were 4.4, 4.3, and 4.3 seconds, respectively. No significant main effect or interaction was found of the factor *navigation interface*.

Hyperlink-rotation

No effect was found for *hyperlink-rotation* ($F_{2,28}<1$). The means for the different hyperlink-rotations of 0, 90, and 180 degrees were 4.6, 4.5, and 4.6 seconds, respectively. No significant main effect or interaction was found of the factor *navigation interface*.

Relative object-bearing

There is a significant effect of the orientation of an object relative to the body ($F_{2,28}=6.66$, $p<0.004$, two-way ANOVA) with an advantage for locations to the front (front=4.0 seconds, side=4.3 seconds, and back=4.6 seconds; Fig. 5.5). A post-hoc Tukey test ($p<0.05$) shows that only the difference in pointing time between the front and the back is significant. No significant effects were found of the factor *navigation interface*.

Object

In Room 1 and 3 significant main effects are found of the factor *object* ($F_{4,56}=4.47$, $p<0.003$; $F_{4,56}=4.51$, $p<0.003$). A post-hoc Tukey test shows significant differences between object five (4.8 seconds) on the one hand and objects one and two (3.7 and 4.0 seconds) on the other hand in Room 1. Furthermore, a significant difference is found between object thirteen (3.2 seconds) on the one hand and objects twelve and fifteen (4.2 and 4.1 seconds) on the other hand, in Room 3. No significant differences between objects were found in Room 2. No significant main effect or interaction was found of the factor *navigation interface*.

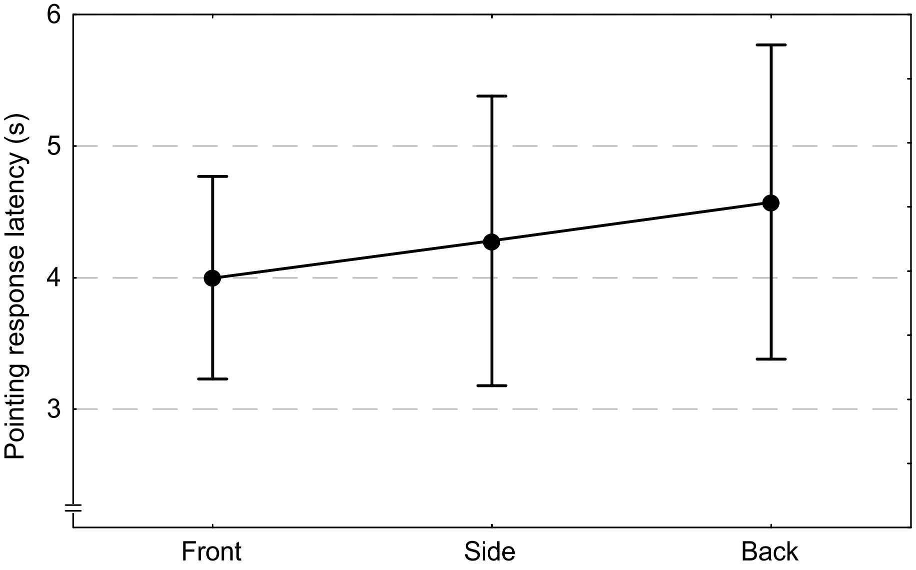


Figure 5.5: Effect of relative-object-bearing. Mean pointing response latencies for objects located to the front, side, or back averaged over participants and navigation group. Whiskers indicate plus and minus one standard deviation, corresponding to variation between participants.

5.2.3 Discussion

Spatial knowledge acquisition

Contrary to expectation, learning the VE with an immersive navigation interface is not faster than with a non-immersive navigation interface. However, the immersive navigation interface shows a trend towards faster learning. The lack of effect might have been caused by the fact that our immersive navigation condition provides only natural navigation for rotations, whereas for translations an input-device was used. It may also be that the abundantly available cues for visual recognition overruled any influence of path integration during VE learning.

Participants who play computer games frequently learn the VE substantially faster than those who don't. Nowadays many computer games require good navigation and spatial orientation skills in three-dimensional graphical environments. This result might suggest that such game experience provides adequate training for spatial orientation in VE. However, the argument can also be turned around, to say that people with good spatial orientation abilities have more success at playing games and therefore like playing computer games more. What is most likely is that both explanations are valid.

There was no effect of immersion on spatial updating performance with the hyperlink displacements. Because no advantages of immersion were found during learning, it is hard to draw any conclusions from this lack of effect.

Spatial updating

Disorientation caused by hyperlinks

The results of the baseline test confirm that participants are temporarily disoriented after a discontinuous displacement and time is needed to update the internal representation to make it correspond to the new orientation. The results of pointing latency show us that there is a significant increase in pointing latency of 0.4 seconds after being moved discontinuously to a new orientation before pointing. This extra time is needed to recognise the new position and to update the internal representation. To compare, in Experiment 6 an extra 0.6 seconds was needed to start moving to a target after discontinuous rotation if cognitive anticipation was excluded. Bowman et al. (1997) found an average increase in search time for targets of 1.2 seconds whenever discontinuous displacement to a location is compared with continuous movement.

The significant and sometimes quite large effect of object (up to one second) found in our study suggests that a substantial part of our response latency is due to cognitive effort involved in remembering object locations.

Does path integration interfere with spatial updating after a hyperlink?

The effect of hyperlink-rotation was tested to show the possible interference of the incorrect location information provided by path integration in the case of hyperlinks. Our results show no effect or trend of an effect of hyperlink-rotation on pointing latency. To compare, Farrell and Robertson (1998) found latency differences of over 1.5 seconds if a movement, and therefore path integration, has to be ignored. In the literature on mental rotation, latency differences are found of up to two seconds if a displacement has to be imagined in the absence of movement stimuli (Boer, 1991; Presson & Montello, 1994; and Rieser, 1989). The main difference between these experiments and our experiment is that in our case visual recognition is involved in spatial updating, which apparently decimates the role of the position information provided by path integration.

A plausible explanation is that path integration is simply *reset* after the hyperlink by visual recognition without incurring any additional cost. With continuous displacement, some form of reset of path integration is needed, at least from time to time, to compensate for error accumulation. It is plausible that such a reset that must be used frequently in everyday locomotion, is performed with little effort. This means that path integration and visual recognition are not mutually independent processes, competing for their contribution to spatial updating. Reasoning thus, path integration is subordinate to visual recognition.

Does the direction of the hyperlink destination influence visual recognition?

No indication was found to support the view that hyperlinks with destinations in the direction of the original viewing direction lead to faster recognition. As a check, the effect of relative object-bearing was measured to confirm the results on which the hypothesis was based. The experiment does indeed confirm the existence of an effect of relative object-bearing. The 0.6 seconds difference between front and back pointing that were found in this experiment are comparable in magnitude with the approximately 0.4 seconds found by both Sholl (1987), and by Easton and Sholl (1995). The fact that no advantage was found for hyperlinks is which the destination is faced is puzzling. Maybe, because participants knew the hyperlink-direction beforehand they were able to focus their spatial attention on the displacement direction even before reaching the start-point. In fact, participants may have performed a mental rotation just before the hyperlink, allowing them to always face the hyperlink destination mentally. This might have obscured the effect that was expected. Indeed, Boer (1991) showed that the effect of relative object-bearing also applies to the imagined heading after mental rotation. In the current experiment there was no sure way of knowing whether participants performed this mental rotation. To clarify this matter, an experiment would have to be performed for hyperlinks in which the hyperlink-direction is not known beforehand.

Note that the hyperlinks were tested only after a mental representation was built by the participants on the basis of free exploration and continuous visual feedback. These results provide no information about how hyperlinks may affect the learning of an environment.

To conclude, no advantage was found of using immersive navigation during continuous movement nor was any disadvantage found of using immersive navigation during discontinuous displacement. There was no evidence to support the notion that the incorrect information provided by path integration interferes with recovery from disorientation after a hyperlink. This suggests that the role of path integration is subordinate to the role of visual recognition in spatial updating.
In the experiment, the possibility for participants to anticipate the spatial destination of the hyperlink was deliberately excluded. The next experiment focuses on the role of cognitive anticipation of the destination of hyperlinks.

5.3 Experiment 9: Cognitive anticipation to hyperlinks

5.3.1 Method

Participants

Twenty participants took part in the experiment, twelve males and eight females. The participants, drawn from different schools and universities, ranged in age from 17 to 23 years, and had normal or corrected to normal eyesight. Participants gave written consent and were paid a fixed amount for their participation.

Task

The experiment consisted of different phases in which either a part of the layout of the VE had to be learned or disorientation was tested following a hyperlink displacement in the learned layout. While learning, participants were instructed to learn the layout of objects in the VE **as fast as possible**. They could move freely in the VE using the navigation interface that was assigned to them. At fixed time-intervals while learning, participants' knowledge was tested by letting them point at objects **as accurately as possible**. Pointing at objects was chosen as a task because this requires that the participants have knowledge of both the environment and their current location. Once the VE had been learned, the disorientation caused by hyperlinks was tested by letting participants point at the objects **as fast as possible** and measuring response latency. Participants had to point at objects after different types of continuous and discontinuous displacement into one of the rooms had occured. The different displacements allow different combinations of the path integration process, the recognition process, and the cognitive anticipation process. By administering the pointing task directly after a displacement, the effectiveness of the spatial updating processes can be measured.

Design and analysis

The factor *navigation interface* (immersive versus non-immersive) was varied between participants. Participants were assigned randomly to either the immersive navigation group or the non-immersive navigation group with the restriction that an equal number of males and females were assigned to each group and the amount of experience with computer games was balanced between both groups.

The second part of the experiment performed by the immersive and the non-immersive group was identical. A total of ten spatial updating tests was administered in which participants had to point at one of the previously learned objects after a specific type of displacement. The displacements differed as far as the availability for spatial updating of cognitive anticipation, visual recognition, and path integration was concerned. Furthermore the tests differed depending on whether or not there was a double-task (Fig. 5.6).

There are three types of displacement which are: continuous movement, hyperlink with recognition but without cognitive anticipation, and hyperlink with cognitive anticipation but without recognition. These three are all executed with and without a double task, making the first six conditions. The four remaining conditions are all executed with a double task. Two more conditions enable recognition alongside of cognitive anticipation. One of these shows a viewpoint consistent with cognitive anticipation and the other shows an inconsistent viewpoint corresponding to a different location from the one anticipated. Participants were warned about this and were told to ignore recognition. The last but one condition shows the destination numbers but only for a short fixed duration after which participants are displaced. The last condition also shows numbers for the same short duration before the hyperlink but now these are different numbers without any relation to the destinations. This last condition serves as a control for the last but one condition.

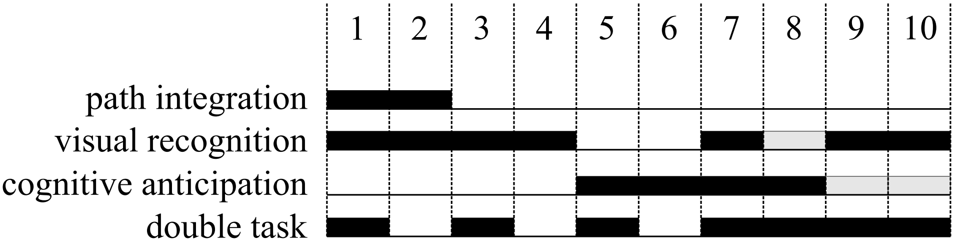


Figure 5.6: Conditions for the 10 tests. A dark bar indicates that the corresponding process can be used to determine the current location when pointing at a target object or indicates the presence of the double task. The grey bars have the following meaning: in Test 8 visual recognition should be ignored because a discrepant viewpoint is shown; in Test 9 anticipation is only possible for a short duration (0.4 s); in Test 10 anticipation is not possible but a distractor is shown for a short duration (0.4 s).

To counteract any effects of fatigue or progressive learning, the order of the ten tests was balanced with a randomised digram-balanced latin-square (Wagenaar, 1969). One participant from the non-immersive and one participant from the immersive group was assigned to each of the 10 different test sequences (Appendix A4).

Each test consisted of 18 pointing trials. The nine different end-points in the three rooms all appeared two times in each test in random order. The objects at which participants had to point were chosen randomly with two restrictions. Firstly, each of the 15 objects appeared at least once during a test. Secondly, the relative object-bearings were balanced so that in each test participants had to point six times at an object located to the front, six times at an object located to the side, and six times at an object located behind. Before administering each test of 18 trials, two extra practice trials were given to participants so that they could to familiarise themselves with the specific test.

The three objects that had to be remembered in the double task were always one from each room and they never contained an object at which participants had to point in the same trial.

Virtual world and stimulus conditions

Virtual world

The virtual world was identical to the world used in the previous experiment (Fig. 5.3).

Object pointer

The appearance of the object pointer was identical to that of the previous experiment (Fig. 5.3). However, to reduce variability in pointing duration the possible attitudes of the dial were reduced to eight discrete angles (0°, 45°, 90°, etc.). The discrete dial prevents participants from spending time on fine-tuning the exact location of the dial, whereas they are only required to point roughly at the target object (within 45°). The orientation of the dial at the start of the trial was randomised, but the dial was never pointing in the correct direction to start with.

Navigation interface

To explore the virtual world, participants used either an immersive navigation interface or a non-immersive navigation interface, identical to that of the previous experiment.

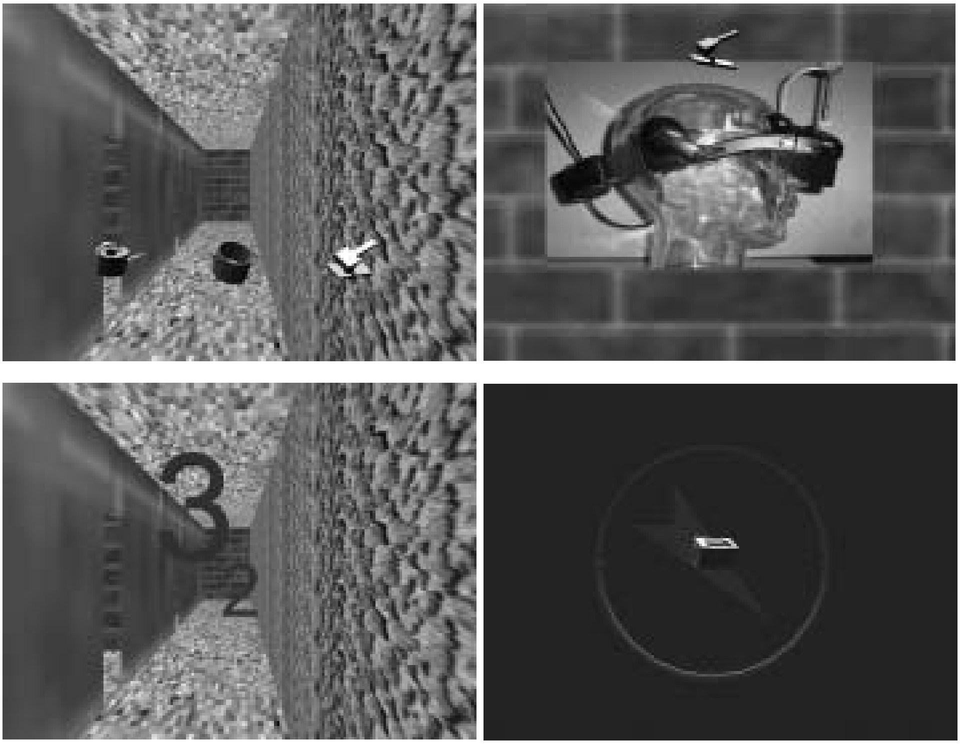


Figure 5.7: Typical views. top-left: presentation of the three objects that have to be remembered for the double task; top-right: presentation of the cue object for the double task at the end of a trial just after pointing; bottom-left: presentation of the room number (large 3) and direction number (small 2) to allow anticipation to the destination of the coming hyperlink; bottom right: pointing to the toaster after a hyperlink with anticipation but with no recognition.

Cognitive anticipation indication.
To enable cognitive anticipation to a hyperlink, a room and a direction number was presented to participants (Fig. 5.7) corresponding to the numbering of rooms and directions in Figure 5.2. This number could be used to determine the exact destination of the hyperlink.

Double task.
The double task consisted of three objects that had to be remembered. The three objects that had to be remembered were shown at the start of a trial on the lower part of the screen for a short duration of 0.8 seconds (Fig. 5.7). This short duration makes it harder to remember all the objects. After displacement and pointing, one of the objects was again shown at the top part of the screen (Fig. 5.7).

Apparatus
The apparatus is identical to that used in the previous experiment.

Procedure
Before the experiment started, a questionnaire was administered on general orientation ability and computer experience. The experiment consisted of three phases: training of the input-device, spatial knowledge acquisition; and spatial updating tests. The total duration of the experiment was between three and four hours. Participants could rest after training, after learning the layout, and after each four tests in phase three.

In **phase one**, participants were trained in navigation with the interface and in using the object pointer, in an identical way to in the previous experiment. The *baseline disorientation test* that was carried out in the previous experiment was not repeated in this experiment because it had already produced significant results.

In **phase two**, the duration of spatial knowledge acquisition was measured in an identical way as in the previous experiment.

In **phase three**, the ten test conditions were administered to investigate the role of path integration recognition and cognitive anticipation in spatial updating. In each trial, participants started in the hallway outside Room 0 facing the direction of the other three rooms.

During all phases of the experiment in which participants had to point at a target object, the twenty objects that were learned were removed from the scene.

The sequence of events during a trial depends on the spatial updating condition. For a discontinuous displacement condition with cognitive anticipation and a double task the sequence is as follows (see also Fig. 5.8): 1) The participant presses any button to indicate readiness for the trial. The participant is next moved with a hyperlink to the start-position in the hallway 2) one second pause; 3) presentation of the double task, the three objects are shown on the bottom part of the screen for 0.8 seconds; 4) the double task objects are removed and after a pause, the room number and the direction number of the destination are shown; 5) The participant presses any button to indicate sufficient cognitive anticipation of the destination. The participant is then directly displaced to the destination location where the numbers are removed and the object pointer immediately appears, with

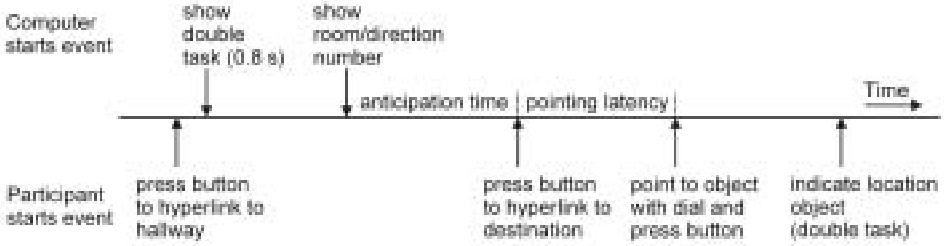


Figure 5.8: Procedure for a trial in a discontinuous condition with a double task.

in front of it the target object; 6) The participant points at the object using the pointer and presses a button; 7) The pointer and the object are removed and the cue objects for the double task response appears in the middle at the top of the screen; 8) The participant presses Button 1, 2, 3 or 4 to indicate the original position of presentation of the cue object; 9) pause until participant indicates readiness for next trial by means of a button press.

For a continuous displacement condition with double task the sequence is as follows: 1), 2) and 3) are identical to the previous sequence; 4) the double task objects are removed and an animation of continuous movement to the end location starts; 5) arrival at the end location, the object pointer immediately appears, with in front of it the target object; 6), 7), 8) and 9) are identical to the previous sequence.

Scoring

For the VE learning phase the dependent measure is again the number of blocks of 90 seconds exploration, needed before participants reached the criterion of doing eighteen pointing trials all with errors less than 45 degrees.

For the pointing tests three dependent measures are used. As in the previous experiment the average pointing latency and the percentage of pointing errors larger than 45 degrees in a test are used. In the tests with a double task, the percentage of error on the double task responses is used. In the tests with self-paced cognitive anticipation, the average cognitive anticipation duration is used. The time needed to give a response to the double task is not used. The reason for this is that the response for the double task involved pressing one of four buttons which had to be blindly searched for. A pilot study had shown that this searching for the right button caused much variability in response latency. Therefore in this case the response latency is not an adequate measure for mental processing time.

After each test condition, participants were asked to give a rating of their mental effort on a nine point scale ranging from *totally no mental effort* up to *extremely high mental effort*. A scale like this can only be used to compare conditions within participants. The scale cannot be used as an absolute measure of mental load.

The data from each participant were cleaned. Pointing latencies smaller than 0.5 seconds were removed. All data points in which the participant did not move the pointer for object-bearing were removed.

5.3.2 Results

Three participants were not able to finish the experiment and were replaced. Of the dropouts, two participants in the immersive group fell ill and one participant in the non-immersive group was unable to learn the layout of objects within two hours.

Spatial knowledge acquisition

The results of the virtual world learning phase show a highly significant effect of *navigation interface* ($F_{1,18}$=18.67, $p<0.001$). The immersive group, with an average of 5.1 learning blocks (std.dev. between participants of 0.71), showed an advantage over the non-immersive group that had an average of 7.9 learning blocks (std.dev. between participants of 1.91). Figure 5.9 shows the learning curves for both groups.

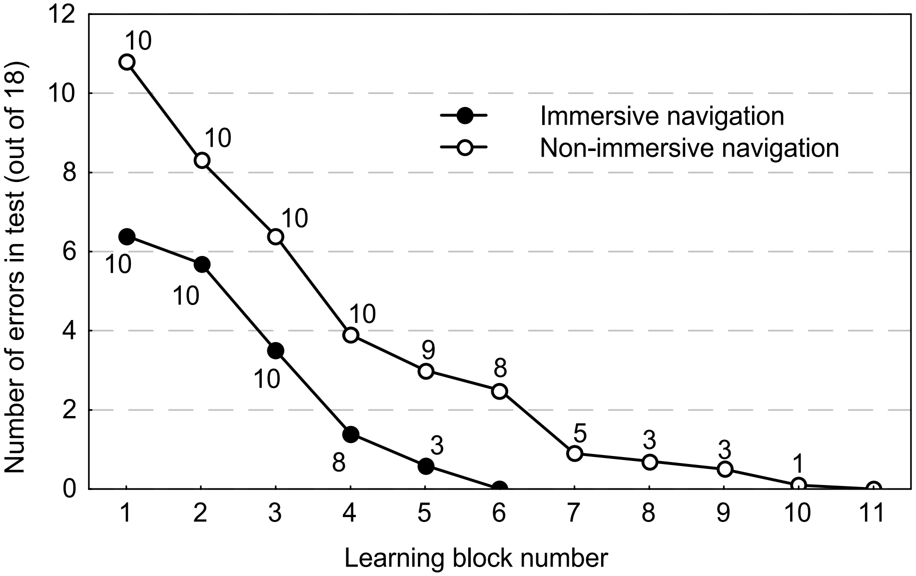


Figure 5.9: Average number of errors in the tests after each learning block for the immersive navigation group and the non-immersive navigation group. At each datapoint the number of participant that still need to reach the training criterion after the corresponding learning block has been indicated. To calculate the average, the number of errors of participants who have already reached the training criterion is taken to be zero.

Pointing latency

To reduce capitalisation on change, the overall effect of the factor test on the pointing latency is firstly analysed with a two-way ANOVA *navigation interface* (2 levels) x *test* (10 levels). Analysis showed a highly significant main effect of *test* ($F_{9,162}$=23.69, p<0.001) and no effect of *navigation interface* or interaction with this factor. The average pointing latencies for all ten tests are given in Figure 5.10. Because *navigation interface* has no effect on pointing latency this factor is not included in subsequent analyses of effects on point latency.

Given the significant overall effect of *test* the results will be analysed in more detail, firstly for Test 1 to 6 involving the three basic displacements with and without a double task. Secondly, the addition of visual recognition to cognitive anticipation will be analysed by

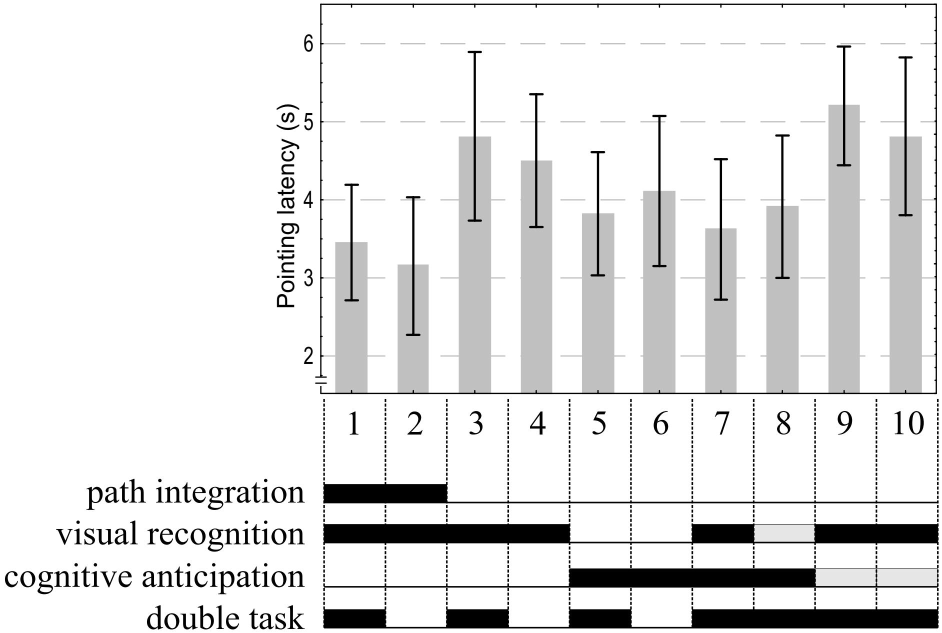


Figure 5.10: Average pointing latency for the ten different test conditions. Whiskers indicate plus and minus one standard deviation corresponding to differences between participants. The dark bars underneath the graph indicate which of the spatial updating processes can be used to determine the current location when pointing at a target object, or to indicate the presence of the double task. The grey bars mean the following: in Test 8 visual recognition should be ignored because a discrepant viewpoint is shown; in Test 9 anticipation is only possible for a short duration (0.4s); in Test 10 anticipation is not possible but a distractor is shown for a short duration (0.4s).

comparing Test 7 with Test 5 to see whether visual recognition still improves performance even if cognitive anticipation is already possible. Furthermore, Test 7 will be compared to Test 1 to see if this potential performance improvement leads to the same performance as with continuous navigation. Thirdly, the influence of a discrepant visual stimulus will be tested by comparing Test 8 with Test 5. Lastly, the effect of the short presentation of destination indications will be analysed by comparing Test 9 with Test3. If this last comparison proves to be significant, Test 10 will be compared to Test 9 to see if these participants are only distracted by the short presentation or if they actually process the room and direction numbers.

The results of pointing latency for Test 1, 2, 3, 4, 5, and 6, are analysed with a two- way ANOVA *double task* (2 levels: present/not-present) x *spatial updating condition* (3 levels: path integration and recognition/ recognition/ cognitive anticipation). A highly significant effect was found of *spatial updating condition* ($F_{2,38}=34.62$, $p<0.001$) and a there was a weak interaction between *spatial updating condition* and *double task* ($F_{2,38}=4.28$, $p<0.02$). On the one hand pointing latencies are reduced for Test 1 and Test 3 if the double task is removed in Test 2 and Test 4, respectively, whereas on the other hand for Tests 5, removal of the double task in Test 6 increases the pointing latency. However, a post-hoc comparison of this interaction with a Tukey test does not show significant differences between Test 1 and 2, nor between Test3 and 4, or between Test 5 and 6.

A post-hoc Tukey test on the main effect of *spatial updating condition* shows a clear difference between all three conditions. The lowest response latencies are found for the two continuous conditions in which path integration is possible (3.3 seconds), followed by the two discontinuous conditions with only self-paced cognitive anticipation (4.0 seconds), and last of all the two discontinuous conditions without cognitive anticipation (4.7 seconds).

The mean of Test 7 (3.6 seconds) with recognition and cognitive anticipation lies between the results of the test with continuous movement (Test 1: 3.4 seconds) and the test with discontinuous displacement and only cognitive anticipation (Test 5: 3.8 seconds). Planned comparison shows, however, that these differences are not significant (Test 1 versus Test 7: $F_{1,19}=0.97$, $p<0.3$; and Test 5 versus Test 7: $F_{1,19}=2.18$, $p<0.2$).

Showing a discrepant viewpoint while pointing at the destination after cognitive anticipation does not increase pointing latency if compared to merely cognitive anticipation with no recognition. (Test 8: 3.9 seconds compared with Test 5: 3.8 seconds).

A short (0.4 seconds) display of the room number and the direction number before the hyperlink (Test 9: 4.8 seconds) shows no measurable difference from the condition in which this short display is not present (Test 3: 4.8 seconds). Given this lack of result it is useless to analyse the results of Test 10 which was a control condition for Test 9.

Pointing errors

A two-way ANOVA *navigation interface* (2 levels) x *test* (10 levels) shows a significant effect of *test* ($F_{9,162}=2.68$, $p<0.006$). The mean error percentages for the ten test conditions were 10%, 9%, 11%, 15%, 15%, 16%, 11%, 18%, 12%, and 15%, respectively. However, a post-hoc Tukey test shows that only the difference between the two lowest errors in Test 1 and 2 on the one hand and the highest error in Test 8 on the other hand are significant.

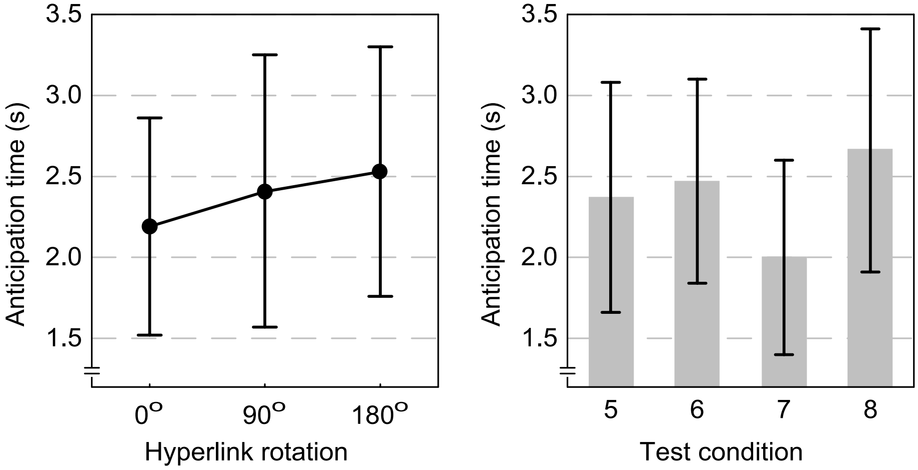


Figure 5.11: Average anticipation time for the three different hyperlink-rotations (averaged over tests) and for Test 7, 8, 9, and 10 (averaged over hyperlink-rotation). In Test 5 and 6 recognition of the own position after the hyperlink is made impossible by putting a black screen behind the object pointer. In Test 7 participants could also rely on recognition after the hyperlink. In Test 8 a discrepant view is present. Whiskers indicate plus and minus one standard deviation corresponding to differences between participants.

Cognitive anticipation duration

The cognitive anticipation duration is analysed with a three-way ANOVA *navigation interface* (2 levels: immersive and non-immersive) x *test* (4 levels: Test 5,6,7, and 8) x *hyperlink-rotation* (3 levels: 0°, 90°, and 180°). Significant main effects are found of *hyperlink-rotation* ($F_{3,36}$=7.94, p<0.001) and of *test* ($F_{3,54}$=5.35, p<0.003).

The time that participants take to anticipate to the discontinuous displacement is larger for bigger hyperlink-rotations (Fig. 5.11). A post-hoc Tukey test shows that the difference between hyperlink-rotations of 90°, and 180° is not significant but that the other differences are significant. In the four different test conditions with cognitive anticipation, the least time is taken to anticipate in the case that recognition can also be used after the displacement in Test 7 (Fig. 5.11). The longest anticipation durations were found in the cases where a discrepant viewpoint is shown after displacement in Test 8. The results of the two tests going on only anticipation (Test 5 and 6) lie between these two values. A post-hoc Tukey test shows that only the difference between the lowest value and the highest two values is significant.

Cognitive anticipation duration was also analysed with a one-way ANOVA for the factor *hyperlink-distance* (3 levels: Room 1, 2, and 3). A significant main effect is found ($F_{2,38}$=3.31, $p<0.05$). Looking at the means for the three rooms (2.3, 2.5, and 2.5 seconds), participants are slightly faster when it comes to anticipating a displacement to the most nearby room in comparison with cognitive anticipation to the two other rooms that are further away.

Double task error

Analyses of the double task error with a two-way ANOVA *navigation interface* (2 levels) x *test* (10 levels) show no significant effects. Only a trend is present ($F_{1,18}$=3.98, $p<0.06$) which suggests that participants in the immersive navigation group might perform better at the double task (14% error) than the participants in the non-immersive group (25% error).

Mental load

A two-way ANOVA *navigation interface* (2 levels) x *test* (10 levels) shows a highly significant effect of *test* ($F_{9,162}$=6.74, $p<0.001$). The mental load ratings for the ten tests are 5.4, 3.7, 5.2, 4.0, 5.2, 4.5, 5.1, 5.8, 5.5, and 5.4.

The lowest mental loads are found in the tests without a double task (Test 2, 4, and 6). The highest load is found in the test with the discrepant viewpoint after cognitive anticipation. A post-hoc Tukey test shows that the difference in rating between Test 2 and Test 4 on the one hand and all other tests on the other hand, except for Test 6 is significant. Test 6 is significantly different from Test 8.

5.3.3 Discussion

Spatial knowledge acquisition

As expected, the immersive navigation interface supports faster spatial learning than the non-immersive navigation interface. The difference between the two interfaces lies in the quality of path integration that is low for the non-immersive interface (Chapter 2). Our results confirm the suggestions made by Schenk (1998), by O'Keefe and Nadel (1978, p94), and by McNaughton et al. (1995) to the effect that path integration helps to encode the location of objects.

Still, the result is surprising because others (Grant & Magee, 1998; and Ruddle et al., 1999) have not found any effects when using direction estimates as a criterion as in the last experiment. An important difference between the present study and the two other studies mentioned is, the scale of the virtual environment. In the present study, a small environment is used in which the topology of the rooms is easily learned but the exact location of an object can only be learned by encoding its metrical location. The large-scale layouts that were used by Ruddle et al. and by Grant and Magee are topologically more complex and there pointing depends less on precise metrical knowledge. The angles in their layouts were primarily 90Ñ. If a participant assumes all path segments to be of equal length, pointing at a distant target will be quite accurate if the topology of the connecting path is known correctly.

As can be seen from prior results, path integration is biased, at least in the case of rotations (Péruch et al., 1997; Chapter 2). The errors accumulate with longer pathways or with more route segments (Klatzky et al. 1990; Loomis et al., 1993; and Sholl, 1989). This would suggest that path integration information could only help to encode the relative location of two objects close to each other. Immersive navigation might be more important for spatial learning on a local level in small layouts than for learning the topology of large layouts.

Once the environment has been completely learned, there is no difference in the speed or accuracy of the pointing task that requires the use of this Cognitive Map (CM) that was built during exploration. This suggests that ultimately there is no difference in the quality of a CM, which is built with either the immersive or the non-immersive interface. This lack of difference might depend, however, on the stage of knowledge acquisition. In the last experiment a very strict training criterion was used ensuring that *complete* knowledge of the environment was obtained. If a fixed short training duration were to be used, the learning curves show (Fig. 5.9) that differences in pointing accuracy would be found. Only a limited set of fifteen objects was used. When using a VE as an interface for the access of information a lot more objects can be expected. This may make the difference in performance between the two navigation interfaces even more pronounced.

Spatial updating

As mentioned in the introduction, path integration, visual recognition, and cognitive anticipation are all involved in spatial updating or in other words the determination of the internal representation of the current location. How these three parallel processes cooperate is still largely unknown. One question that was posed in the introduction was how effective cognitive anticipation is in comparison with the other modes of spatial updating. Furthermore, does cognitive anticipation to a hyperlink follow similar dynamics as mental displacements that are not actually executed? Also, is there a difference in the mental effort needed for the spatial updating processes and are processes triggered automatically or are they under cognitive control?

The effectiveness of cognitive anticipation

The effectiveness of cognitive anticipation is most clearly shown by the pointing latencies after a displacement. As expected, without cognitive anticipation to the hyperlink destination, pointing latencies are largest, which reconfirms the results established in Chapter 4. However, the pointing latencies with cognitive anticipation are still substantially larger than with continuous displacement where path integration and visual recognition can be used. As could be expected in this case, the result of pointing latencies for spatial updating for the condition with also recognition besides cognitive anticipation lies in between the results for the only cognitive anticipation condition and the continuous movement condition. These last differences are, however, too small to be significant.

Pointing errors were generally low and little difference was found between test conditions. The only difference to be found was between the lowest errors for the two continuous conditions and the highest error for the cognitive anticipation condition with a discrepant stimulus for recognition. This last difference indicates that some confusion may have occurred due to the discrepant posters. Apparently, participants do not trade off accuracy for speed. The lack of difference in errors between different types of displacement might suggest that the best way to transport ourselves through a VE would only depend on the time spent in transportation including the cognitive anticipation. However, only displacement in a very well known layout was investigated. If the participants have incomplete knowledge of a layout, it becomes apparent from the different learning curves found in the experiments that differences in error occur that depend on the navigation interface. It would be useful to investigate how fast the continuous movement can be made before the advantage for spatial updating breaks down. In such an investigation the update rate would be an important variable.

The dynamics of cognitive anticipation

Our results showed that the cognitive anticipation time for a hyperlink increases with larger hyperlink-rotations and larger hyperlink-distances. This is very similar to the effects found in mental displacement (Boer, 1991; Easton & Sholl, 1995; Presson & Montello, 1994; and Rieser, 1989). If the location and orientation after a hyperlink are closer to the original location and orientation, less time is needed for cognitive anticipation. However, the magnitude of the effects that were found is substantially smaller than the effects that are te be found in the literature on mental rotation. A difference was found of 0.2 seconds between a hyperlink-rotation of 0Ñ and 90Ñ. The difference between these angles in terms of pointing latency for mental rotation is well over one second (Boer,1991; Easton & Sholl, 1995; and Rieser, 1989). A possible explanation may be found in the interpretation of the mental rotation task. Wraga et al. (2000) found that there is a substantial difference between instructing participants to rotate themselves mentally or to rotate the surrounding array mentally, although the resulting relative object locations are the same. The reaction time differences between 0Ñ and 90Ñ for self-rotation are 0.2 seconds, whereas for object array rotations this difference is around 0.8 seconds. In our experiment, participants were instructed to prepare themselves for the hyperlink as fast as possible in order to be able to point as fast as possible after displacement. This instruction does not require participants to mentally simulate the movement. Another difference with our study was that besides having a rotation the participant also were translated at the same time.

These results suggest that when incorporating hyperlinks in an interface design, hyperlinks without rotations are to be used. However, there may be many more factors that are yet unknown, which may have an influence on the dynamics of cognitive anticipation.

Significant differences were found in cognitive anticipation times in the four conditions that included cognitive anticipation. The lowest cognitive anticipation time to be found is if visual recognition of the current location is also possible after the hyperlink. The highest cognitive anticipation time to be found is when participants know that a discrepant viewpoint will be shown. It seems that participants brace themselves to not use visual recognition.

If the room and orientation number are only shown for a short duration there is no time for anticipation and participants simply seem to ignore the numbers at no additional cost. Performance with short presentation times is in all respects equal to performance in the condition in which only visual recognition after the hyperlink is possible.

Mental load of the spatial updating processes
A double task was included in most of the conditions to increase the mental load. The double task should have interfered with performance in the spatial updating conditions that already have a high workload. Although the double task presumably increases the mental load, the results show no conclusive evidence of interference between the double task and spatial updating. A non significant trend is visible, indicating an increase in workload if one goes from continuous navigation to discontinuous navigation with only recognition and to discontinuous navigation with cognitive anticipation. The double task errors show no differences between the test conditions. No effects were found on the pointing latency for the different test conditions relating to the presence or absence of the double task. A weak interaction is found between the factor double task and the three main spatial updating conditions, but this interaction is very weak. The interaction cannot be interpreted because none of the underlying differences between conditions with or without a double-task are significant.

Automatism of recognition or cognitive anticipation
The results of the condition in which recognition should be ignored are not clear-cut. If recognition is automatic, performance decrements were expected if an inconsistent viewpoint would be shown after a hyperlink. There was an effect on the percentage of pointing error but only in comparison with the two continuous conditions. No effect was found on pointing latency. A trend is found for an increase in mental effort. A significant increase was found in cognitive anticipation duration. Together, these effects indicate that participants experience at least some difficulty with suppressing visual recognition in the condition where that was required.

Cognitive anticipation is not automatically executed but only by deliberate intent. However, the process could be impossible to stop once it has been consciously activated. This would be shown by a performance decrement in the condition where the cognitive anticipation numbers are only shown briefly to participants. No evidence was found of this possibility.

5.4 Conclusions

The results show that immersive navigation has advantages over non-immersive navigation when it comes to acquiring spatial knowledge in a VE. The locations of objects are remembered faster because the additional feedback from the body with an immersive interface helps to encode the relative locations of objects. Once the VE is learned, no differences are found in spatial updating performance between the two navigation interfaces. This indicates that the ultimate quality of the acquired CM does not differ in both conditions.

The fastest way of achieving displacement through a VE is by using a hyperlink, but the using of hyperlinks causes temporary disorientation, especially if cognitive anticipation is not possible. Contrary to what was expected, no evidence was found to support the notion that the type of hyperlink affects visual recognition efficiency after displacement.

Cognitive anticipation of the destination of the hyperlink rules out disorientation after the hyperlink to a great extent. However, even with cognitive anticipation some time is needed to recover from a discontinuous displacement, which is not the case when arriving at a place after a continuous movement.

The time needed for cognitive anticipation depends on the characteristics of the hyperlink. Hyperlinks in which the viewing direction in the VE is changed take longer to anticipate to. A hyperlink over larger distances also requires more time for cognitive anticipation.

6 General discussion

6.1 Summary of the results

Three main research questions were raised in Chapter 1:

Q1. When compared to non-immersive navigation, does immersive navigation improve the quality of path integration?
Q2. When compared to non-immersive navigation, does immersive navigation improve the acquisition of a cognitive map?
Q3. Does discontinuous displacement affect spatial orientation?

Q1 was investigated in Chapter 2 and Chapter 3. In Chapter 2 path integration was investigated with different combinations of visual, vestibular, and kinesthetic feedback. The results showed that navigation with an immersive interface that provides additional kinesthetic feedback will lead to a higher quality of path integration than navigation with a non-immersive interface that provides only visual feedback. There is not only a difference in the biases of path integration with both the interfaces but also a difference in performance variation. In addition to these objective results, the subjective evaluation of the different conditions shows a clear preference for the immersive navigation interface.

Even with the immersive navigation interface, a considerable accumulation of errors in path integration is seen. This means that visual recognition is needed to ensure accurate spatial updating. The low quality of path integration suggests that visual recognition might be more dominant in spatial updating and that path integration is only used on a very local scale or during brief intervals when recognition is not possible.

Chapter 3 investigated whether training can compensate for the degraded path integration performance that was found with a non-immersive interface. The results showed that the biases in purely visual path integration can be greatly reduced by providing participants with knowledge of the results or by enabling them to use visual recognition. Although the performance improvement is still present when testing the one day later, the trend in the data suggests that no long term improvement of performance is to be found with training. Furthermore, even directly after training the variability in performance remains, showing that the visual stimulus used provides poor information for path integration.

Reports on the strategies that were used, together with several verifications of the stimulus material, suggest that most people tend to follow the displacement of specific objects rather than to estimate velocity and integrate this over the course of time.

When path integration training is provided to a non-immersive group before exploring a new environment (Chapter 4), no improvement can be found in spatial knowledge acquisition although an improvement was expected. In the experiment in question, the participants' task was to find an object. Indeed some participants reported that often when they had found the object in question, they did not know their own location, which indicates a poor quality of spatial updating.

These results confirm Hypothesis 1: An interface that allows immersive navigation can improve the quality of path integration when compared to an interface that only allows non-immersive navigation.

Q2 was investigated in the first parts of all four of the experiments that were reported in Chapter 4 and Chapter 5. Together these experiments show that immersive navigation improves the acquisition of a CM when compared with non-immersive navigation. Still, it was only in the last of the four experiments that the effect was significant. The other three experiments showed trends pointing in the same direction.

The advantage of immersive navigation for spatial knowledge acquisition is hard to show because large individual differences exist between participants' abilities to acquire a CM. These differences are partly caused by differences in innate spatial orientation abilities. Furthermore, these differences derive from differences in training or experience with the interface. This was shown in the improving of learning performance with repeated experience with the VE interface in Chapter 4. Another source of variability between participants in spatial knowledge acquisition is the amount of experience with computer games. This influence of computer game experience may be an indication of an effect of training in spatial orientation tasks through experience, but it may also point to higher innate abilities, leading to an enjoyment in playing computer games.

Once a CM has been acquired, no differences in performance are found between immersive and non-immersive navigation. Having a CM of the environment allows participants to rely on recognition. The possibility of having visual recognition probably decimates the influence of path integration. Furthermore, even if the visual information for recognition were poor, cognitive anticipation can be used to determine the location purely on the basis of the combination of the CM and knowledge about the movements to be executed.

These results confirm Hypothesis 2: Immersive navigation can improve the acquisition of a cognitive map.

Q3 was divided into three parts:

Q3a Does discontinuous navigation impair the acquisition of a cognitive map?

Q3b Is cognitive anticipation able to compensate for the disruption of spatial updating that is caused by discontinuous displacement?

Q3c Does the type of discontinuous displacement have any effect on spatial updating?

Question Q3a was investigated in the first experiment in Chapter 4. As expected the results showed that navigation by discontinuous displacement degrades spatial knowledge acquisition. The only way for participants to integrate the visual information that is offered by the non-overlapping views is by using cognitive anticipation. Knowledge about the spatial relationship between two subsequent views before and after displacement is essential for establishing the current location. Without this knowledge it would be impossible to relate the location of two objects seen in different views, at least in a not yet known VE.

These results confirm Hypothesis 3a: Discontinuous navigation can impair the acquisition of a cognitive map.

Question Q3b was investigated by comparing different combinations of the three spatial updating processes. If no possibility for cognitive anticipation is provided, then discontinuous displacement causes disorientation that takes time to recover from and sometimes leads to errors in decisions depending on correct spatial updating. If complete information about the destination of a discontinuous displacement is provided beforehand, then disorientation after the hyperlink will be greatly reduced. However, spatial updating with continuous navigation is still superior. These results show that although the availability of cognitive anticipation is essential for reducing disorientation by hyperlinks, cognitive anticipation may not be sufficient to completely compensate for the loss of all the perceptual information that is available during continuous navigation.

The effectiveness of the mental processes involved in spatial updating and the mental effort that is needed to execute the process is related to the degree to which the processes are executed automatically. Automatism is strongly related to experience. Since path integration and visual recognition are used almost continuously in everyday life we may expect these processes to be automatic. Cognitive anticipation of a discontinuous displacement does not occur in everyday life and is therefore assumed without further investigation to not be automatic.

For path integration without visual feedback, others have convincingly shown that spatial updating is executed automatically and cannot even be ignored without creating a considerable performance decrement (Farrell and Robertson, 1998; and Farrell and Thomson, 1998). For path integration on the basis of a visual stimulus only this might be different. The experiments in Chapter 2 and 3 have not only shown that visual path integration is unreliable, but debriefing also showed that most participants have to follow objects attentively in order to be able to determine their displacement. Furthermore, training visual path integration does not improve spatial knowledge acquisition, which might indicate that participants do not even use visual path integration when exploring a VE. Although all the separate evidence is circumstantial, when put together these findings do suggest that visual path integration is not automatic, at least not with the artificial stimuli that are presented in an HMD.

The automatism of visual recognition was explored by investigating whether a discrepant visual stimulus could be ignored without cost. Several trends were found that suggest, if taken together, that visual recognition cannot be ignored without cost.

These results confirm Hypothesis 3b: Cognitive anticipation can reduce disorientation caused by discontinuous displacement.

Question Q3c was investigated for path integration, visual recognition and cognitive anticipation in the second parts of the experiments in Chapter 5.

With path integration, it was expected that the incorrect information provided by this process after discontinuous displacement might interfere with spatial updating. There was not even a hint of evidence that the incorrect information provided by the path integration process would interfere with spatial updating. Apparently the information coming from the path integration process is ignored. However, this does not necessarily mean that ignoring path integration information is effortless. There might still be some effort involved in the cognitive control process of ignoring the information. Part of the increase in response latency that we find after a hyperlink might be caused by the necessity to reset the path integration process.

For visual recognition, it was speculated that recovery from discontinuous displacement would be easier if the destination of the displacement was to lie more in the viewing direction at the start of the hyperlink. No hint of evidence was found to indicate that the type of hyperlink affects the efficiency of visual recognition after a hyperlink. However, a shortcoming of the experiment set-up was that the direction of displacement was known beforehand. Just before the hyperlink, participants could have performed a mental rotation that would allow them to mentally occupy a viewing direction that faces the hyperlink destination independent of the actual viewing direction. However, this is unlikely because participants don't report any difference in difficulty between the starting points when asked afterward, and mental rotation requires very deliberate effort.

With cognitive anticipation, a highly reliable relation was found between the type of hyperlink and the time needed for cognitive anticipation. Cognitive anticipation requires more time if the viewing direction in the VE is changed during the displacement. A small effect of the distance of displacement was also found.

These results confirm Hypothesis 3c: The type of discontinuous displacement can have an effect on spatial updating.

6.2 Implications of the results

As was described in the introduction, Chapter 1, the operators that supervise complex processes need to retrieve information about the state of the system that is being controlled from the interface. The operator has to gather and integrate information that is distributed across the information space as presented by the interface. In order to gather information, the operator navigates both by continuous and by discontinuous displacements. A growing interest was observed for the application of VE technology as an alternative interface to support operators. A distinction was made between immersive VE technology and non-immersive VE technology.

In this thesis, the possible consequences of the choice of interface technology were investigated for continuous as well as discontinuous navigation tasks. To investigate these consequences, a series of nine experiments was carried out using exemplary VEs that were designed to isolate the different effects expected.

The VEs that were used in the experiments were of limited size and had only a limited number of objects. When supervising complex processes, the amount of information that is needed will generally be much larger than used in the experiments. For this reason, the results that were found in the experiments do not apply quantitatively to VE interfaces designed to support operator. Qualitatively, however, the same effects as were found in the experiments may be expected in a real VE operator interface. Increasing the amounts of information in a VE would make it more difficult to find information. This probably would increase the quantitative differences found in this thesis.

The choice of interface technology

The experiments have shown clear advantages of immersive navigation over non-immersive navigation. Performances in both path integration and spatial knowledge acquisition are improved with immersive navigation. However, a distinction should be made between novice and more experienced users of the interface. After the cognitive map of the layout of the VE was acquired, no advantages were found of immersive navigation over non-immersive navigation. Furthermore, repeated use of the non-immersive interface improves performance in both path-integration and spatial knowledge acquisition.

These results suggest that immersive navigation might only be beneficial for application domains in which every time new spatial layouts have to be learned or in domains where the primary users are novices. For instance, in training firemen to teach them the layout of new buildings, or for architectural walkthroughs to evaluate new building designs. For supervisory control applications, the advantages of immersive navigation will only have an effect during familiarisation with the interface in which the layout of information has to be learned. After knowledge of the layout is acquired, no continuing benefit of immersive navigation should be expected.

When having to make a choice in interface technology, the advantages that were found of immersive navigation have to be considered as well as the disadvantages. The current immersive VE technology causes eyestrain, headaches or even nausea to many users. These problems will have to be solved first if widespread use of the technology is to be enabled.

Furthermore, immersive technology is still considerably more expensive than non-immersive technology. To give an indication: a state-of-the-art head-tracker for six degrees-of-freedom costs around % 14.000, whereas a six degrees-of-freedom input-device sells at around % 500. An HMD costs anywhere between % 10.000 for a reasonable quality product and over % 100.000 for top quality products, whereas a conventional computer screen of a reasonable size can be bought for around % 1.000. The graphic computers that would be needed to render a 3D scene are the same for an immersive and a non-immersive interface.

Graphic computer performance is still increasing dramatically, whereas prices remain to drop. However, the peripherals needed for immersive VE like head-trackers and HMDs are only used by a select company. This means that production series are small, prices are high, and development is slow because there are only a few manufacturers investing.

The choice for the type of displacement

Discontinuous displacement disrupts spatial updating but it also has the advantage of being fast. When exploring a new environment discontinuous movement should not be allowed because this hinders the acquisition of a CM. Once the environment is known this recommendation will change. If time is not a critical factor, continuous movement should be clearly preferred. If fast displacement is essential then discontinuous displacement should be preferred.

The disorienting effects of discontinuous displacement can be greatly reduced by allowing for cognitive anticipation. The interface designer must make sure that information is provided about the destination of a discontinuous displacement.

The type of discontinuous displacements has an effect on the time needed for anticipation. Discontinuous displacements that involve a rotation take more time to anticipate and should, if possible, be avoided. However, other effects may also govern spatial updating during discontinuous displacements and so more research is needed.

Besides being valid for navigation in supervisory control interfaces, these recommendation apply to navigation in 3D data-sets in general like, for instance, navigation in 3D medical images such as MRI and CT scans, or navigation through 3D environments on the World Wide Web.

6.3 Recommendation for future research

In Chapter 1 a Framework for the Investigation of Navigation and Disorientation (FIND) was presented (Fig. 1.2). This model helps to understand in what ways the interface can influence task performance, and it may also be useful for others who are interested in spatial orientation in VE. Especially the inclusion of cognitive anticipation as a separate process in FIND proved to be an important factor in better understanding spatial updating during discontinuous displacements.

In retrospect, a minor modification should be made to improve FIND. Based on the poor performance that was found for visual path integration, it seems better to delete the visual input of the path integration process from FIND altogether.

Based on the results of our experiments some recommendations can be be made for the design or the selection of participants in future experiments. It is general practise in spatial orientation research to balance experimental designs for gender. However, rarely can balancing for computer game experience be found. In Experiment 8 a large effect was found of computer game experience on acquisition of spatial knowledge in a VE. For future research, it is advisable to equally divide participants according to their computer game experience over the experimental groups.

Similarly, the results of the experiments in Chapter 3 and Chapter 4 showed that prior experience with navigation in a VE can have a large effect on performance in spatial orientation tasks. Visual path integration can easily be calibrated and spatial knowledge acquisition improves with increasing VE experience. Experimenters should be wary of these results and should act either to balance their designs, or to carefully select their participants.

The research presented in this thesis constitutes only a first step towards discovering the consequences of allowing discontinuous displacements in a VE. A lot more needs to be known in order to be able to provide complete guidelines for interface design. In the next paragraphs, recommendations are made for research directly related to the research presented in this thesis.

The continuum between continuous and discontinuous navigation

As discussed in the introduction, continuous and discontinuous navigation are only two extremes in a sliding scale. A clear advantage of discontinuous navigation is the reduction in travel time. However, the results of this thesis show that this may disrupt spatial updating. Time is needed to anticipate the displacement beforehand, and even with such preparation, time is still needed to recover from the disorientation caused by the discontinuous displacement.

A solution that benefits from both worlds might be that of increasing the speed of the continuous displacement. Because movement at very high speeds would become impossible to steer, the movements would have to be automated, which means that the viewpoint displacement would be observed passively.

The question is, how much can the speed be increased before the advantages of continuous movement start to break down and spatial updating performance drops? To answer this question one must consider both the speed of movement and the update rate of the image generator. Together these two variables determine the amount of visual overlap between the two consecutive frames shown. Performance might drop because the speed of movement simply becomes too high or because the amount of overlap in two subsequent frames becomes too low.

More dependencies of cognitive anticipation time on the type of hyperlink

As explained in Chapter 5, the spatial characteristics of a hyperlink can be defined using the hyperlink rotation, the hyperlink distance and the hyperlink direction. The last experiment showed that the cognitive anticipation time depends on hyperlink rotation and on hyperlink distance. Hyperlink direction was not investigated to limit the size of the experiment.

Given these results it is reasonable to think that more dependencies may exist between the type of hyperlink and the cognitive anticipation time. For instance, the results produced by Boer (1991), Sholl (1987), and Easton and Sholl (1995) showed that spatial locations are more easily remembered if one faces the spatial location in question. Hyperlinks that displace the viewpoint in the direction being faced might be easier to cope with than hyperlinks that displace the viewpoint to the side or back. The given spatial characteristics need to be expanded if the displacements are also to include a change in altitude or rotations about one of the horizontal axes.

The non-spatial characteristics of a hyperlink may also affect cognitive anticipation. Investigating the dependencies between the type of hyperlink and the cognitive anticipation time is important because it provides guidelines for the design of 3D interfaces.

The relation between the type of CM and spatial updating during hyperlinks

In this thesis the relation between the type of hyperlink and the spatial updating processes was explored. However, both visual recognition and cognitive anticipation also depend on the use of the CM. Therefore, the properties of this CM may well affect spatial updating during hyperlinks. Two such properties are the viewpoint dependency of the CM and the hierarchical nature of the CM.

Viewpoint dependency means that information about a spatial layout is stored in a preferred orientation (Diwadkar & McNamara, 1997; Roskos-Ewoldsen, McNamara, Shelton, & Carr, 1998; Shelton, & McNamara, 1997). The effect of this viewpoint dependency is that pointing to unseen objects in the environment is faster and more accurate if one stands at the dominant viewpoint. Viewpoint dependency especially occurs with layouts that are only learned from limited viewpoints. Even if three or four different viewpoints are presented during spatial knowledge acquisition, the CM is not viewpoint invariant but represents multiple viewpoints (Shelton & McNamara, 1997). However, with more extensive exploration the CM has no preferred orientation (Presson & Hazelrigg, 1984; Sholl, 1987).

The consequence of having a viewpoint-dependent CM is that when imagining this layout this is always done initially from one preferred viewpoint. Diwadkar and McNamara (1997) have even shown that the time needed to recognise a scene is shorter when presented from the dominant viewpoint. Hyperlinks that displace an operator to the dominant viewpoint may require less effort to cope with.

Evidence exists to show that spatial information is stored in the CM hierarchically, as if the space was divided into regions (Hirtle & Jonides, 1985; McNamara, Hardy, & Hirtle, 1989). The hierarchical organisation leads to distortions, meaning, for instance, that the distance between two locations in different regions is overestimated when compared to equal distances between two locations in the same region. A hyperlink that crosses the boundaries of the hierarchical regions in our CM may be harder to cope with than hyperlinks that remain within one region of our CM.

Appendix A: within-participants designs

Design of Experiment 1 (Chapter 2)

For an explanation of the conditions (I to V) see Section 2.2.1.

Participant	Stimulus condition				
1	I	V	III	II	IV
2					
3	IV	III	II	I	V
4					
5	II	IV	V	III	I
6					
7	III	I	IV	V	II
8					
9	V	II	I	IV	III
10					

Design of Experiment 2 (Chapter 2)

For an explanation of the conditions (VI to VIII) see Section 2.3.1.

Participant	Stimulus condition		
1	VI	VII	VIII
2			
3	VI	VIII	VII
4			
5	VII	VI	VIII
6			
7	VII	VIII	VI
8			
9	VIII	VI	VII
10			
11	VIII	VII	VI
12			

Design of Experiment 6 (Chapter 4)

A=immersive continuous, B= non-immersive continuous,
C=immersive discontinuous, and D=non-immersive discontinuous.
For a detailed description of these conditions see Section 4.2.1.

	Navigation condition (A..D). World (1..4)			
Participant	Session 1	Session 2	Session 3	Session 4
1	D.1	B.2	A.3	C.4
2	D.2	B.4	A.1	C.3
3	D.3	B.1	A.4	C.2
4	D.4	B.3	A.2	C.1
5	B.1	C.2	D.3	A.4
6	B.2	C.4	D.1	A.3
7	B.3	C.1	D.4	A.2
8	B.4	C.3	D.2	A.1
9	A.1	D.2	C.3	B.4
10	A.2	D.4	C.1	B.3
11	A.3	D.1	C.4	B.2
12	A.4	D.3	C.2	B.1
13	C.1	A.2	B.3	D.4
14	C.2	A.4	B.1	D.3
15	C.3	A.1	B.4	D.2
16	C.4	A.3	B.2	D.1

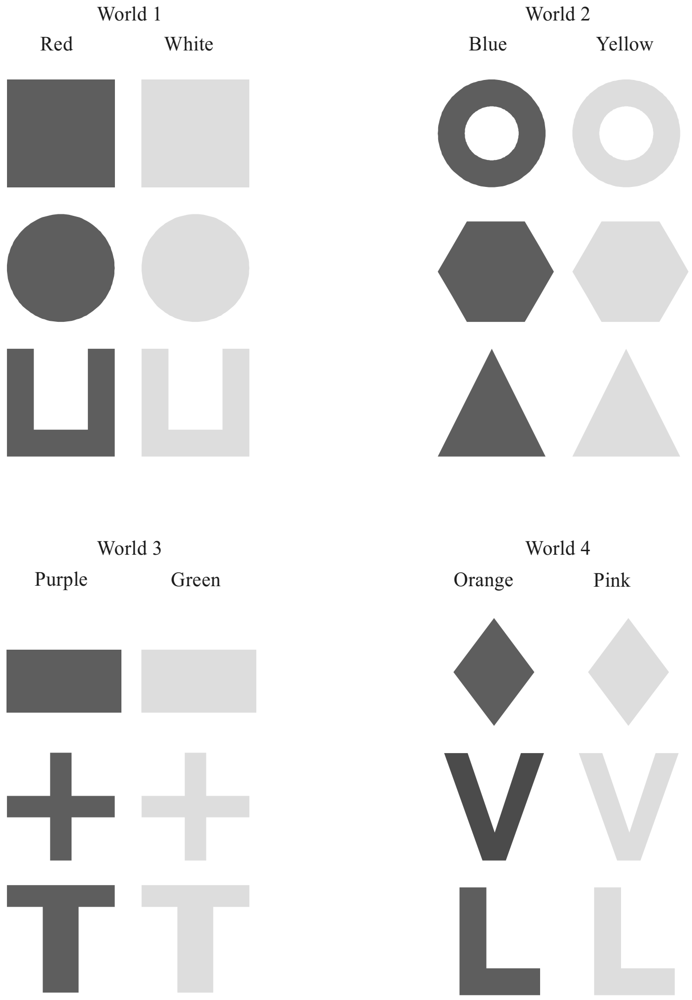
Design of Experiment 9 (Chapter 5)

For a detailed description of the conditions (1…10) see Section 5.3.1.

Participant	Navigation	Test condition										
1	Immersive	Practise	1	7	8	5	6	4	2	9	3	10
2	Non immersive											
3	Immersive		7	5	1	4	8	9	6	10	2	3
4	Non immersive											
5	Immersive		8	1	6	7	2	5	3	4	10	9
6	Non immersive											
7	Immersive		5	4	7	9	1	10	8	3	6	2
8	Non immersive											
9	Immersive		6	8	2	1	3	7	10	5	9	4
10	Non immersive											
11	Immersive		4	9	5	10	7	3	1	2	8	6
12	Non immersive											
13	Immersive		2	6	3	8	10	1	9	7	4	5
14	Non immersive											
15	Immersive		9	10	4	3	5	2	7	6	1	8
16	Non immersive											
17	Immersive		3	2	10	6	9	8	4	1	5	7
18	Non immersive											
19	Immersive		10	3	9	2	4	6	5	8	7	1
20	Non immersive											

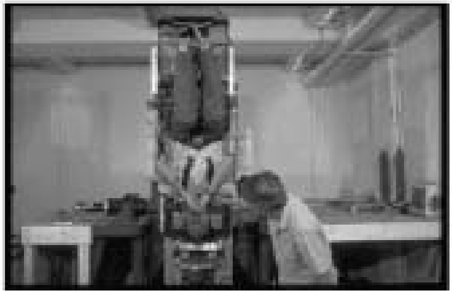
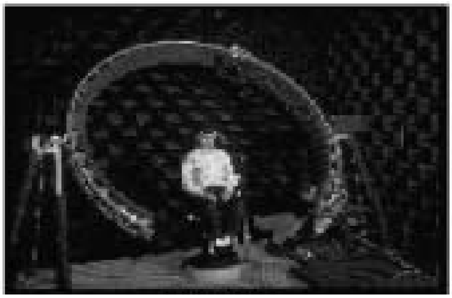
Appendix B: virtual objects

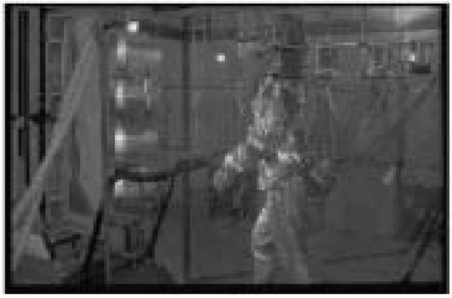
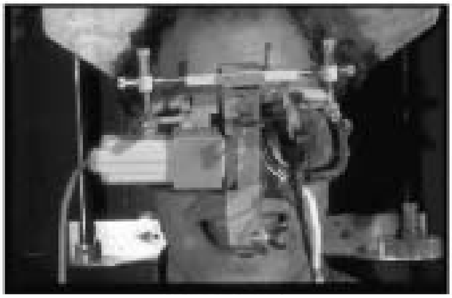
Set of vitual objects that were used in Experiment 6 and 7 (Section 4.2.1/4.3.1)



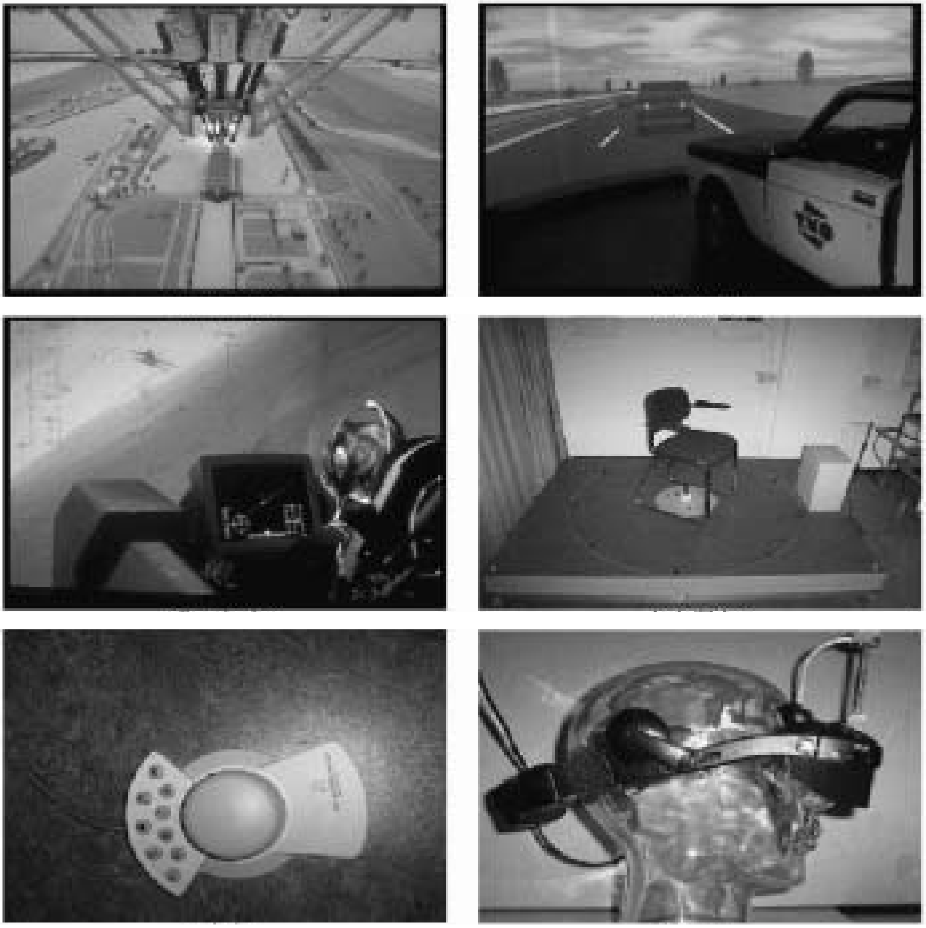
Set of virtual objects that were used in Experiment 8 and 9
(Section 5.2.1/Section 5.3.1)

Set of posters that were used in Experiment 8 and 9
(Section 5.2.1/Section 5.3.1)

Set of posters that were used in Experiment 8 and 9 (Section 5.2.1/Section 5.3.1)



References

Akatsuka, Y., & Bekey, G.A. (1998). Compensation for end to end delays in a VR system. *Proceedings of the IEEE Virtual Reality Annual International Symposium (VRAIS'98), Atlanta, GA*, pp.156-159.

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

Bakker, N.H., Passenier, P.O., & Werkhoven P.J. (1998). Spatial orientation in virtual environments: isolating the roles of head-slaved vision and continuous visual feedback. *Proceedings of the 17th European Annual Conference on Human Decision Making and Manual Control, Valenciennes, France*, pp.187-196.

Bakker, N.H., Passenier, P.O, & Werkhoven, P.J. (submitted) The effects of immersion and discontinuous displacement on spatial orientation in virtual environments. *Human Factors.*

Bakker, N.H., Passenier, P.O, Werkhoven, P.J., Stassen, H.G., & Wieringa, P.A. (submitted) Navigation in virtual environments. IFAC conference 2001.

Bakker N.H., Werkhoven P.J., & Passenier P.O. (1997). Orientation in virtual environments. *Proceedings of the 16th European Annual Conference on Human Decision Making and Manual Control, Kassel, Germany*, pp.208-209.

Bakker, N.H., Werkhoven, P.J., & Passenier, P.O. (1998). Aiding orientation performance in virtual environments with proprioceptive feedback. *Proceedings of the IEEE Virtual Reality Annual International Symposium (VRAIS'98), Atlanta, GA*, pp.28-33.

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

Bakker, N.H., Werkhoven, P.J., & Passenier, P.O (2000). Recovery from Disorientation by 3D Hyperlinks. *(Report TNO-TM-00-D007), TNO Human Factors, Soesterberg, The Netherlands*, pp.19.

Bakker, N.H., Werkhoven, P.J., & Passenier, P.O. (in press). Calibrating visual path integration in VE. *Presence.*

Barfield, W., Hendrix, C., & Bystrom, K.-E. (1999). Effects of stereopsis and head tracking on performance using desktop virtual environment displays. *Presence*, 8(2), 237-240.

Barnes, G.R., & Sommerville, G.P. (1978). Visual target acquisition and tracking performance using a helmet-mounted sight. *Aviation, Space, and Environmental Medicine*, April, 565-572.

Beintema, J. (2000). Self-motion perception from optic flow and rotation signals. *Unpublished Ph.D Thesis, Erasmus University, Rotterdam, The Netherlands*.

Bertin, R.J.V., Israël, I., & Lappe, M. (in press). Perception of two-dimensional, simulated ego-motion trajectories from optic flow. *Vision Research*.

Billinghurst, M., Bowskill J., Dyer N., & Morphett J. (1998). An evaluation of wearable information spaces. *Proceedings of the IEEE Virtual Reality Annual International Symposium (VRAIS'98), Atlanta, GA*, pp.20-27.

Billingsley, P.A. (1982). Navigation through hierarchical menu structures: does it help to have a map. *Proceedings of the Human Factors Society 26th Annual Meeting, Seattle, WA*, pp.103-107.

Bles, W. (1981). Stepping around: circular vection and coriolis effects. *In J. Long & A.Baddeley (Eds.), Attention and Performance IX, Hillsdale, N.J.*.

Bles, W., Vianney De Jong, J.M.B., & De Wit, G. (1984). Somatosensory compensation for loss of labyrinthine function. *Acta Otolaryngol*, 97, 213-221.

Bliss, J.P., Tidwell, P.D., & Guest, M.A. (1997). The effectiveness of virtual reality for administering spatial navigation training to firefighters. *Presence*, 6(1), 73-86.

Boer, L.C. (1991). Mental rotation in perspective problems. *Acta Psychologica*, 76, 1-9.

Boman D.K. (1995). International survey: virtual-environment research. *IEEE Computer*, June, 57-65.

Bowman, D.A., Koller D., & Hodges, L.F. (1997). Travel in immersive virtual environments: an evaluation of viewpoint motion control techniques. *Proceedings of the IEEE Virtual Reality Annual International Symposium (VRAIS'97), Albuquerque, New Mexico*, pp.45-52.

Brandt, T., Wist, E.R., & Dichans, J. (1975). Foreground and background in dynamic spatial orientation. *Perception & Psychophysics*, 17(5), 497-503.

Breedveld, P. (1996). The design of a man-machine interface for a space manipulator. *Unpublished Ph.D Thesis, Delft University of Technology, The Netherlands*.

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*, 7(2), 168-178.

Chen Y., Bedell, H.E., & Frishman, L.J. (1998). The precision of velocity discrimination across spatial frequency. *Perception & Psychophysics*, 60(8), 1329-1336.

Christou, C., & Bulthoff, (U-umlaut) H.H. (1998). Differences between active-explorers and passive-observers in virtual scene recognition. *(Technical Report No.62), Max-Planck-Institut fur (u-umlaut) biologische Kybernetik, Tubingen, Germany*.

Cohen, R. (Ed.) (1985). The development of spatial cognition. *Lawrence Erlbaum Associates Inc. Publishers, Hillsdale, New Jersey*.

Darken R.P., & Sibert J.L. (1996). Navigating large virtual spaces. *International Journal of Human-Computer Interaction*, 8(1), 49-71.

Diwadkar, V.A., & McNamara, T.P. (1997). Viewpoint Dependence in Scene Recognition. *Psychological Science*, 8(4), 302-307.

De Vries, S.C., & Padmos, P. (1997). Remotely controlled flying aided by a head-slaved camera and HMD. *(Report TNO-TM-97-B024), TNO-Human Factors, Soesterberg, The Netherlands*, p.19.

Dennehy, M.T., Nesbitt, D.W., & Sumey, R.A. (1994). Real-time three-dimensional graphics display for antiair warfare command and control. *John Hopkins APL Technical Digest*, 15(2), 110-119.

Easton, R.D., & Sholl, M.J. (1995). Object-array structure, frames of reference, and retrieval of spatial knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(2), 483-500.

Edgar, G.K, Pope, J.C.D, & Craig, I.R. (1994). Visual accommodation problems with head-up and helmet-mounted displays. *Displays*, 15(2), 68-75.

Edwards, K.L., Buckle, J.W., Doherty, M.J., Lee, L.J., Pratty, A.C., & White, J.F. (1997). An operationally-applicable objective method for the analysis and evaluation of the flights of helicopter mission task elements during field-of-view trials. *Proceedings of the 53rd Annual Forum of the AHS, Virginia Beach, VA*.

Ellis S.R. (Ed.) (1993). Pictorial communication in virtual and real environments. *Taylor & Francis Ltd., London*.

Elm, W.C., & Woods, D.D. (1985). Getting lost: a case study in interface design. *Proceedings of the Human Factors Society 29th Annual Meeting, Baltimore, Maryland*, pp.927-931.

Farrell, M.J., & Robertson, I.H. (1998). Mental rotation and the automatic updating of body-centered spatial relationships. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(1), 227-233.

Farrell, M.J., & Thomson, J.A. (1998). Automatic spatial updating during locomotion without vision. *The Quarterly Journal of Experimental Psychology*, 51A(3), 637-654.

Foreman, N., & Gillett, R. (Eds.) (1997). Handbook of spatial research paradigms and methodologies - Vol.1 spatial cognition in the child and adult. *Psychology Press, East Sussex, U.K.*.

Foreman, N., & Gillett, R. (Eds.) (1998). Handbook of spatial research paradigms and methodologies - Vol.2 clinical and comparative studies. *Psychology Press, East Sussex, U.K.*.

Frank, L.H., Casali, J.G., & Wierwille, W.W. (1988). Effects of visual display and motion system delays on operator performance and uneasiness in a driving simulator. *Human Factors*, 30(2), 201-217.

Gaillard, A.W.K. (1996). Stress produktiviteit en gezondheid (Stress productivity and health). *Uitgeverij Niewezijds, Amsterdam, The Netherlands*.

Gaunet, F., Martinez, J.-L., & Thinus-Blanc, C. (1997). Early-blind subjects' spatial representation of manipulatory space: exploratory strategies and reaction to change. *Perception*, 26, 345-366.

Gibson (1979). The ecological approach to visual perception. *Houghton Mifflin Company, Boston, MA*.

Gibson, J.J. (1950). The perception of the visual world. *The Riverside Press, Cambridge, MA*.

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

Griffin, M.J. (1990). Handbook of human vibration. *Academic Press Limited, London*.

Hasebe, H., Oyamada, H., Ukai, K., Toda, H., & Bando, T. (1996). Changes in oculomotor functions before and after loading of a 3-D vissually-guided task by using a head-mounted display. Ergonomics, 39(11), 1330-1343.

Healy S. (Ed.) (1998). Spatial representation in animals. *Oxford University Press, Oxford.*

Hirtle, S.C., & Jonides, J. (1985). Evidence of hierarchies in cognitive maps. *Memory & Cognition*, 13(3), 208-217.

Howard, I.P. & Templeton, W.B. (1966). Human spatial orientation. *John Wiley & Sons, London*,.

Howarth, P.A, & Finch, M. (1999). The nauseogenicity of two methods of navigating within a virtual environment. *Applied Ergonomics*, 30(1), 39-45.

Iggo, A. (Ed.) (1973). Handbook of sensory physiology. *Springer-Verlag, Berlin.*

Israël, I., Sievering, D., & Koenig, E. (1995). Self-rotation estimate about the vertical axis. *Acta Otolaryngol*, 115, 3-8.

Israël, I., Bronstein, A.M., Kanayama, R., Faldon, M., & Gresty, M.A. (1996). Visual and vestibular factors influencing vestibular "navigation". *Experimental Brain Research*, 112, 411-419.

Ivanenko, Y.P., Grasso, R., Israël I., & Berthoz, A. (1997). The contribution of otoliths and semicircular canals to the perception of two-dimensional passive whole-body motion in humans. *Journal of Physiology*, 502.1, 223-233.

Iwata, H., & Fujii, T. (1996). Virtual perambulator: a novel interface device for locomotion in virtual environments. *Proceedings of the IEEE Virtual Reality Annual International Symposium (VRAIS'96), Santa Clara, CA*, pp.60-65.

Iwata, H., & Yoshida, Y. (1999). Path reproduction tests using a torus treadmill. *Presence*, 8(6), 587-597.

Jobling, J.T., Mansfield, J.S., Legge, G.E., & Menge, M.R. (1997). Motion parallax: effects of blur, contrast, and field size in normal and low vision. *Perception*, 26, 1529-1538.

Johnson, D.M., & Stewart, J.E. (1999). Use of virtual environments for the acquisition of spatial knowledge: comparison among different visual displays. *Military Psychology*, 11(2), 129-148.

Kalawsky, R.S. (1993). The science of virtual reality and virtual environments. *Addison-Wesley Publishing Company, Wokingham, UK.*

Kappé (1997). Visual information in virtual environments. *Unpublished Ph.D thesis, University of Utrecht, The Netherlands.*

Kaur, K., Sutcliffe, A.G., & Maiden, N.A.M. (1999). Towards a better understanding of usability problems with virtual environments. *Proceedings of the IFIP TC.13 International Conference on Human-Computer Interaction (INTERACT'99), Edinburgh, UK.*, pp.527-535.

Kawara, T., Ohmi, M., & Yoshizawa, T. (1996). Effects on visual functions during tasks of object handling in virtual environment with a head mounted display. *Ergonomics*, 39(11), 1370-1380.

Kenyon, R.V., & Kneller, E.W. (1992). Human performance and field of view. *SID International Symposium Digest of Technical Papers, Santa Ana, CA*, pp.290-293.

Klatzky R.L., Loomis J.M., Golledge R.G., Cicinelli J.G., Doherty S., & Pellegrino J.W. (1990). Acquisition of route and survey knowledge in the absence of vision. *Journal of Motor Behavior*, 22(1), 19-43.

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.

Koenderink, J.J. (1986). Optic flow. *Vision Research*, 26(1), 161-180.

Kolasinski, E.M., Goldberg, S.L., & Hiller J.H. (1995). Simulator sickness in virtual environments. *(Technical Report 1027, Army Project Number 2O262785A791), U.S. Army Research Institute for the Behavioral and Social Sciences, Alexandria, VA*.

Kuipers, B. (1982). The "map in the head" metaphor. *Environment and Behavior*, 14(2), 202-220.

Lampton, D.R., Knerr, B.W., Goldberg, S.L., Bliss, J.P., Moshell, J.M., & Blau, B.S. (1994). The Virtual Environment Performance Assesment Battery (VEPAB): development and evaluation. *Presence*, 3(2), 145-157.

Land, M.F. (1992). Predictable eye-head coordination during driving. *Nature*, 359, 318-320.

Lederman, S.J., Klatzky. R.L., Collins, A., & Wardell, J. (1987). Exploring environments by hand or foot: time-based heuristics for encoding distance in movement space. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(4), 606-614.

Loomis, J.M., Da Silva, J.A., Fujita, N., & Fukusima, S.S. (1992). Visual space perception and visually directed action. *Journal of Experimental Psychology: Human Perception and Performance*, 18(4), 906-921.

Loomis, J.M., Klatzky, R.L., Golledge R.G., Cicinelli, J.G., Pellegrino, J.W., & Fry P.A. (1993). Nonvisual navigation by blind and sighted: assessment of path integration ability. *Journal of Experimental Psychology: General*, 122(1), 73-91.

Mackinlay, J.D., Card S.K., & Robertson, G.G. (1990). A semantic analysis of the design space of input devices. *Human Computer Interaction*, 5, pp.145-190.

Maurer, R., & Séguinot, V. (1995). What is modelling for? A critical review of the models of path integration. *Journal of Theoretical Biology*, 175, 457-475.

May, M. (1996). Cognitive and embodied modes of spatial imagery. *Psychologische Bëitrage*, Band 38, 418-434.

May, M., & Klatzky, R.L. (2000). Path integration while ignoring irrelevant movement. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(1), 169-186.

McConkie, G., & Rudmann, D.S. (1998). Acquiring spatial knowledge from varying fields of view. *Proceedings of 2nd Annual FEDLAB Symposium: Advanced Displays & Interactive Displays, College Park, MD*, pp.49-53.

McCreary, F.A., & Williges, R.C. (1998). Effects of age and field-of-view on spatial learning in an immersive virtual environment. *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting, Chicago*, pp.1491-1495.

McGee, M.K. (1998). Using psychophysics to measure negative side effects of immersive virtual environments. *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting, Chicago*, pp.1501-1505.

McNamara, T.P., Ratcliff, R., & McKoon G. (1984). The mental representation of knowledge acquired from maps. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(4), 723-732.

McNamara, T.P. (1986). Mental representations of spatial relations. *Cognitive Psychology*, 18, 87-121.

McNamara, T.P., Hardy, J.K., & Hirtle, S.C. (1989). Subjective hierarchies in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(2), 211-227.

McNaughton, B.L., Knierim, J.J., & Wilson, M.A. (1995). Vector encoding and the vestibular foundations of spatial cognition: neurophysiological and computational mechanisms. *In The Cognitive Neurosciences, M.S. Gazzaniga (Editor in chief). MIT Press, Cambridge, MA*, pp.585-595.

Mittelstaedt, M.-L., & Mittelstaedt, H. (1996). The influence of otoliths and somatic graviceptors on angular velocity estimation. *Journal of Vestibular Research*, 6(5), 355-366.

Mon-Williams, M., Wann, J.P., & Rushton, S. (1993). Binocular vision in a virtual world: visual deficits following the wearing of a head-mounted display. *Ophthalmic and Physiological Optics*, 13, 387-391.

Mon-Williams, M., & Wann, J.P. (1998). Binocular virtual reality displays: when problems do and don't occur. *Human Factors*, 40(1), 42-49.

Montello D.R. (1991). Spatial orientation and the angularity of urban routes - a field study. *Environment and Behavior*, 23(1), 47-69.

Moseley, M.J., & Griffin, M.J. (1986). Effects of display vibration and whole-body vibration on visual performance. *Ergonomics*, 29(8), 977-983.

Mulder, M. (1999). Cybernetics of tunnel-in-the-sky displays. *Ph.D Thesis, Delft University Press, The Netherlands*.

O'Keefe, J, & Nadel, L. (1978). The hippocampus as a cognitive map. *Oxford University Press, Oxford*.

Onishi, S., Yoshimatsu, H., Kawamura, A., & Ashizaki, K. (1994). An approach to natural vision using a novel head-mounted display. *SID International Symposium Digest of Technical Papers, Santa Ana, CA*, pp.28-31.

Osgood, R.K., & Wells, M.J. (1991). The effect of field-of-view size on performance of a simulated air-to-ground night attack. *Proceedings of Helmet Mounted Displays and Night Vision Goggles, AGARD-CP-517, Neuilly sur Seine, France*, pp.10.1-10.7.

Padmos, P. (1999). Head slaved displays in a driving simulator. *(Report TNO-TM-99B008), TNO-Human Factors, Soesterberg, The Netherlands.*

Paige, G.D., & Seidman, S.H. (1999). Characteristics of the VOR in response to linear acceleration. *In B. Cohen & B.J.M. Hess (Eds), Annals of the New York Academy of Sciences: Otolith Function in Spatial Orientation and Movement. The New York Academy of Sciences, New York*, 871, 123-135.

Palmisano S. (1996). Perceiving self-motion in depth: the role of stereoscopic motion and changing-size cues. *Perception & Psychophysics*, 58(8), 1168-1176.

Péruch, P., Vercher, J.-L., & Gauthier, G.M. (1995). Acquisition of spatial knowledge through visual exploration of simulated environments. *Ecological Psychology*, 7(1), 1-20.

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

Péruch, P. (1999). Transfer of spatial knowledge between virtual and real environments: which information could be transferred. *Abstract, presented at the Workshop on Spatial Cognition in Real and Virtual Environments, Max-Planck-Institut für (u-umlaut) biologische Kybernetik, Tübingen, Germany.*

Peterson, B., Wells, M., Furness, T.A., & Hunt, E. (1998). The effects of the interface on navigation in virtual environments. *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting, Chicago*, pp.1496-1500.

Presson C.C. & Hazelrigg M.D. (1984). Buiding spatial representations through primary and secondary learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(4), 716-722.

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

Rasmussen J. (1986). Information Processing and Human-Machine Interaction. *Elsevier Science Publishing B.V., Amsterdam, The Netherlands.*

Rennison, E., & Strausfeld, L. (1995). The millenium project: constructing a dynamic 3+D virtual environment for exploring geographically, temporally and categorically organized historical information. *Proceedings of the International Conference COSIT, Semmering, Austria*, pp.69-91.

Riecke, B.E. (1998). Untersuchung des menschlichen Navigationsverhaltens andhand von Heimfindeexperimenten in virtuellen Umgebungen.. *Unpublished M.Sc. thesis, Max-Planck-Institut für biologische Kybernetik, Tübingen, Germany*, .

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.

Risch, J., May, R., Thomas, J., & Dowson, S. (1996). Interactive information visualization for exploratory intelligence data analysis. *Proceedings of the IEEE Virtual Reality Anual International Symposium (VRAIS'96), Santa Clara, CA*, pp.230-238.

Roskos-Ewoldsen, B., McNamara, T.P., Shelton, A.L., & Carr, W. (1998). Mental Representations of Large and Small Spatial Layouts Are Orientation Dependent. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(1), 215-226.

Roth, S.F., Chuah, M.C., Kerpedjiev, S., Kolojejchick, J.A., & Lucas, P. (1997). Toward an information visualization workspace: combining multiple means of expression. *Human-Computer Interaction*, 12, 131-185.

Royden, C.S., Banks, M.S., & Crowell, J.A. (1992). The perception of heading during eye movements. *Nature*, 360, 583-585.

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

Ruddle, R.A., Payne, S.J., & Jones D.M. (1998). Navigating large-scale "desk-top" virtual buildings: effects of orientation aids and familiarity. *Presence*, 7(2), 179-192.

Ruddle, R.A., Payne, S.J., & Jones, D.M. (1999). The effects of maps on navigation and search strategies in very-large-scale virtual environments. *Journal of Experimental Psychology: Applied*, 5(1), 54-75.

Rushton, S., Mon-Williams, M., & Wann, J.P. (1994). Binocular vision in a bi-ocular world: new-generation head-mounted displays avoid causing visual deficit. *Displays*, 15(4), 255-260.

Rushton, S.K., & Riddell, P.M. (1999). Developing visual systems and exposure to virtual reality and stereo displays: some concerns and speculations about the demands on accommodation and vergence. *Applied Ergonomics*, 30, 69-78.

Sadalla, E.K., & Montello, D.R. (1989). Remembering changes in direction. *Environment and Behavior*, 21(3), 346-363.

Sandor, P.B., & Leger, A. (1991). Tracking with a restricted field of view: performance and eye-head coordination aspects. *Aviation, Space, and Environmental Medicine*, November, 1026-1031.

Schenk, F. (1997). The Morris maze (is not a maze). In N. Foreman & R. Gillett (Eds.) Handbook of spatial research paradigms and methodologies Vol.2 - clinical and Comparative Studies, Psychology Press, East Sussex, U.K., pp.145-189.

Schryver J.C. (1994). Experimental validation of navigation workload metrics. *Proceedings of the Human Factors Society 38th Annual Meeting, Nashville, Tennessee*, pp.340-344.

Shelton, A.L., & McNamara, T.P. (1997). Multiple Views of Spatial Memory. *Psychonomic Bulletin & Review*, 4(1), 102-106.

Sheridan, T.B. (1988). Task allocation and supervisory control. *In M. Helander (Ed.), Handbook of Human-Computer Interaction, Elsevier Publishing, Amsterdam*, pp.159-173.

Shiwa, S., & Miyasato, T. (1997). A new HMD causing natural accommodative changes. *SID International Symposium Digest of Technical Papers, Santa Ana, CA*, pp.265-268.

Sholl, M.J. (1987). Cognitive maps as orienting schemata. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(4), 615-628.

Sholl, M.J. (1989). The relation between horizontality and rod-and-frame and vestibular navigational performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(1), 110-125.

Smets, G.J.F. (1992). Designing for telepresence: the interdependence of movement and visual perception implemented. *Proceedings of the 5th IFAC / IFIP / IFORS / IEA Symposium on Analysis, Design, and Evaluation of Man-Machine Systems, The Hague, The Netherlands*, 3.1.3.

Smets, G.J.F., & Overbeeke, K.J. (1995). Visual resolution and spatial performance: the trade-off between resolution and interactivity. *Proceedings of the IEEE Virtual Reality Annual International Symposium (VRAIS'95), Research Triangle Park, North Carolina*, pp.67-73.

So, R.H.Y., & Griffin, M.J. (1993). Effect of lags on human performance with head-coupled simulators. *(Report AL/CF-TR-1993-0101), Armstrong Laboratory, Wright-Patterson Air Force Base, Ohio*.

Spillmann, L., & Werner, J.S. (Eds.) (1990). Visual perception - the neurophysiological foundations. *Academic Press Inc., San Diego, CA*.

Stanney, K.M., & Kennedy, R.S. (1997). The psychometrics of cybersickness. *Communications of the ACM*, 40(8), 67-68.

Stassen, H.G., Johannsen, G., & Moray, N.P. (1990). Internal representation, internal model, human performance model and mental workload. *Automatica*, 26(4), 811-820.

Stassen, H.G., & Smets, G.J.F. (1995). Telemanipulation and telepresence. *Proceedings of the 6th Annual IFAC/IFORS/IFIP/SEA Symposium on Man-Machine Systems, Boston, Massachusetts*, pp.13-23.

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

Sugihara, T., & Miyasato, T. (1998). A lightweight 3-D HMD with accommodative compensation. *Proceedings of the Society for Information Display International Symposium, Anaheim, CA*, pp.927-930.

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

Theeuwes, J.L. (1992) Selective attention in the visual field. *Unpublished PhD Thesis, The Vrije Universiteit, Amsterdam, The Netherlands*.

Thorndyke, P.W., & Hayes-Roth, B. (1982). Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology*, 14, 560-589.

Van Erp, J.B.F., & Van den Dobbelsteen J.J. (1998). Head-slaved camera control, time delays, and situation awareness in UAV operation. *(Report TNO-TM-98-A075), TNO-Human Factors, Soesterberg, The Netherlands*.

Van Erp, J.B.F., & Van den Dobbelsteen J.J. (1998). Head-slaved and manual remote camera control with time delays. *(Report TNO-TM-98-A076), TNO-Human Factors, Soesterberg, The Netherlands*.

Van Veen, H.-J., & Riecke, B. (1999). Visual homing to a virtual home. *Abstract, presented at the Workshop on Spatial Cognition in Real and Virtual Environments, Max-Planck-Institut für biologische Kybernetik, Tübingen, Germany*.

Venturino, M., & Kunze, R.J. (1989). Spatial awareness with a helmet-mounted display. *Proceedings of the Human Factors Society 33rd Annual Meeting, Denver, Colorado*, pp.1388-1391.

Vicente, K.J., Hayes, B.C., & Williges, R.C. (1987). Assaying and isolating individual differences in searching a hierarchical file system. *Human Factors*, 29(3), 349-359.

Vicente, K.J., & Rasmussen, J. (1992). Ecological interface design: theoretical foundations. *IEEE Transactions of Systems, Man, and Cybernetics*, 22(4), 589-606.

Vishton, P.M., & Cutting J.E. (1995). Wayfinding, displacements, and mental maps: velocity fields are not typically used to determine one's aimpoint. *Journal of Experimental Psychology: Human Perception and Performance*, 21(5), 978-995.

Voorhorst, F.A. (1998). Affording action - implementing perception-action coupling for endoscopy. *Unpublished Ph.D. Thesis, Delft University of Technology, The Netherlands*.

Wagenaar, W.A. (1969). Note on the construction of digram-balanced latin squares. *Psychological Bulletin*, 72(6), 384-386.

Waller, D., Hunt, E., & Knapp, D. (1998). The transfer of spatial knowledge in virtual environment training. *Presence*, 7(2), 129-143.

Wan, H.S., Touretzky, D.S., & Redish, A.D. (1993). Towards a computational theory of rat navigation. *Proceedings of the 1993 Connectionist Models Summer School, M. Mozer, P. Smolensky, D. Touretzky, J. Elman, and A. Weigend (Eds.), Lawrence Erlbaum Associates, Hillsdale, NJ*, pp.11-19.

Wang, R.F., & Simons, D.J. (1999). Acitive and passive scene recognition across views. *Cognition*, 70, 191-210.

Ware, C., & Osborne, S. (1990). Exploration and virtual camera control in virtual three dimensional environments. *Computer Graphics*, 24(2), 175-183.

Ware, C., & Slipp, L. (1991). Using velocity control to navigate 3D graphical environments: a comparison of three interfaces. *Proceedings of the Human Factors Society 35th Annual Meeting, San Fransisco, CA*, 300-304.

Warren, W.H. Jr., Morris, M.W., & Kalish, M. (1988). Perception of translational heading from optical flow. *Journal of Experimental Psychology: Human Perception and Performance*, 14(4), 646-660.

Warren, W.H. Jr., Blackwell, A.W., Kurtz, K.J., Hatsopoulos, N.G., & Kalish, M.L. (1991). On the sufficiency of the velocity field for perception of heading. *Biological Cybernetics*, 65, 311-320.

Warren, W.H., & Kurtz, K.J. (1992). The role of central and peripheral vision in perceiving the direction of self-motion. *Perception & Psychophysics*, 51(5), 443-454.

Wells, M.J., Peterson, B., & Aten, J. (1997). The virtual motion controller: a sufficient-motion walking simulator. *Proceedings of the IEEE Virtual Reality Annual International Symposium (VRAIS'97), Albuquerque, New Mexico*, pp.1-8.

Wells, M.J., Venturino, M., & Osgood, R.K. (1988). Using target replacement performance to measure spatial awareness in a helmet-mounted simulator. *Proceedings of the Human Factors Society 32nd Annual Meeting, Annaheim, CA*, pp.1429-1433.

Wells, M.J., & Venturino, M. (1990). Performance and head movements using a helmet-mounted display with different sized field-of-view. *Optical Engineering*, 29(8), 870-877.

Wells, M.J., Peterson, B., & Aten, J. (1997). The virtual motion controller: a sufficient-motion walking simulator. *Proceedings of the IEEE Virtual Reality Annual International Symposium (VRAIS'97), Albuquerque, New Mexico*, pp.1-8.

Werkhoven P.J. & Groen J. (1998). Manipulation performance in interactive virtual environments. *Human Factors,* 40(3), 432-442.

Werkhoven, P.J., & Hoekstra W. (1994). Realization of a virtual environment. *(Report: TNO-TM-1994-B-21), Soesterberg, The Netherlands.*

Wickens, C.D. (1992). Engineering psychology and human performance. *Second Edition, HarperCollins Publishers, New York, NY.*

Wickens, C.D., & May, P. (1994). Terrain representation for air traffic control: a comparison of perspective with plan view displays. *(Report ARL-94-10/FAA-94-2) Federal Aviation Administration, Washinton, D.C..*

Wilson, P.N., Foreman, N., & Tlauka, M. (1997). Transfer of spatial information from a virtual to a real environment. *Human Factors*, 39(4), 526-531.

Wist, E.R., Diener, H.C., Dichgans, J., & Brandt, Th. (1975). Perceived distance and the perceived speed of self-motion: linear vs. angular velocity?. *Perception & Psychophysics*, 17(6), 549-554.

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, 413-428.

Witmer, B.G., & Sadowski Jr., W.J. (1998). Nonvisually guided locomotion to a previously viewed target in real and virtual environments. *Human Factors*, 40(3), 478-488.

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

Woodin, M.E., & Allport, A. (1998). Indepent reference frames in human spatial memory: body-centered and environment-centered coding in near and far space. *Memory & Cognition*, 26(6), 1109-1116.

Woods, D.D. (1984). Visual momentum: a concept to improve the cognitive coupling of person and computer. *International Journal of Man-Machine Studies*, 21, 229-244.

Wraga, M., Creem, S.H., & Proffitt, D.R. (2000). Updating displays after imagined object and viewer rotations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(1), 151-168.

Yoshida, A., Rolland, J.P., & Reif, J.H. (1995). Design and applications of a high-resolution insert head-mounted-display. *Proceedings of the IEEE Virtual Reality Anual International Symposium (VRAIS'95), Research Triangle Park, North Carolina*, pp.84-93.

Yungkurth Hooper E. & Coury B.G. (1994). Graphical displays for orientation information. *Human Factors*, 36(1), 62-78.

Zhai, S., & Milgram, P. (1993). Human performance evaluation of isometric and elastic rate controllers in a 6 DOF tracking task. *Proceedings of SPIE 2057, Technical Conference of Telemanipulator Technology and Space Telerobotics, Boston*, pp. 130-141.

Samenvatting

Ruimtelijke Oriëntatie in Virtuele Omgevingen

Operators, die verantwoordelijk zijn voor de supervisie van complexe processen, worden de laatste decennia geconfronteerd met een toenemende hoeveelheid informatie, die nodig is om de processen te beheersen. De huidige grafische interfaces, die gebruikt worden om de informatie te presenteren, bestaan doorgaans uit een beperkt aantal conventionele computerbeeldschermen. Omdat de hoeveelheid informatie te groot is om gelijktijdig weer te geven, worden meerdere presentaties gebruikt waartussen geschakeld moet worden. Het vinden van de juiste informatie in de interface en het behouden van een goed overzicht betekent een aanzienlijke belasting voor de operator.

Recentelijk is er een groeiende interesse voor de toepassing van *virtuele omgevingen* als interface-technologie. Virtuele omgevingen zijn drie-dimensionale, door de computer gegenereerde, beelden die veelal getoond worden in een zogenaamd hoofdgekoppeld display waarin twee kleine beeldschermen vlak voor de beide ogen zijn geplaatst. Wanneer bovendien gebruik gemaakt wordt van sensoren voor het bepalen van de hoofdpositie en de gegenereerde beelden dienovereenkomstig aangepast worden, kan de gebruiker op een natuurlijke manier rondkijken in de virtuele omgeving.

Het gebruik van een virtuele omgeving als interface biedt de mogelijkheid om data in zijn natuurlijk formaat te tonen, bijvoorbeeld de 3D-positie van een vliegtuig ten behoeve van luchtverkeersleiding. Bovendien wordt de interface-ontwerper meer vrijheid geboden bij het organiseren en indelen van de data. Het zoeken van informatie kan echter ook bij gebruik van deze 3D-interfaces een aanzienlijke taakbelasting zijn voor de operator.

Verschillende typen technologie voor virtuele omgevingen kunnen worden onderscheiden en bovendien kunnen verschillende manieren van navigeren mogelijk worden gemaakt. Er moet een keus gemaakt worden tussen de twee hoofdtypen van technologie:

- Een *onderdompelende* interface waarbij rijke sensorische terugkoppeling aanwezig is als de gebruiker zich verplaatst in de virtuele omgeving. Dit wordt mogelijk gemaakt door gebruik te maken van hoofdpositiesensoren gecombineerd met een hoofdgekoppeld display.
- Een *niet-onderdompelende* interface waarbij tijdens het bewegen alleen visuele terugkoppeling aanwezig is. De interface bestaat uit een conventioneel beeldscherm en de beweging door de omgeving wordt indirect bestuurd met een invoerapparaat zoals bijvoorbeeld een computermuis.

Verder moet een keus gemaakt worden tussen twee manieren van verplaatsen:

- *Continue* verplaatsing waarbij het gezichtspunt vloeiend verplaatst wordt door de virtuele omgeving.
- *Discontinue* verplaatsing waarbij het gezichtspunt instantaan verplaatst wordt over een willekeurig grote afstand. Een discontinue verplaatsing heeft als voordeel dat hij zeer snel is. Een nadeel is mogelijk, dat de verplaatsing zelf niet kan worden waargenomen, hetgeen desoriëntatie zou kunnen veroorzaken.

Er is onvoldoende kennis over de invloed van ieder van deze alternatieven op het functioneren van een operator die gebruik maakt van de interface.

Om inzicht te verkrijgen in de consequenties van deze keuze is een raamwerk geformuleerd dat de ruimtelijke oriëntatie van een operator in een virtuele omgeving kwalitatief beschrijft. Om informatie op te zoeken moet een operator beslissen waarheen hij, of zij, zich wil verplaatsen. Om deze beslissing te kunnen maken is kennis nodig van de virtuele omgeving welke opgeslagen is in het ruimtelijk geheugen, de zogenaamde *cognitieve kaart*. Verder moet de operator de *huidige locatie* weten. Deze wordt bepaald door een combinatie van drie processen: *padintegratie,* de registratie van een ruimtelijke verplaatsing op grond van de aanwezige sensorische informatie; *visuele herkenning,* waarbij de locatie direct herkend wordt uit het geboden visuele beeld; *cognitieve anticipatie* de bepaling van de locatie op basis van de cognitieve kaart en kennis van voorgenomen acties, bijvoorbeeld bij een discontinue verplaatsing. Alle processen worden gecontroleerd door een cognitief regelproces dat, afhankelijk van de huidige taak, de beperkt beschikbare aandacht verdeelt over de processen.

Op grond van een literatuuroverzicht zijn drie vragen geformuleerd over de invloed van de verschillende keuzen op het gedrag van operators.

1. Verbetert een onderdompelende interface de kwaliteit van padintegratie door de toegenomen sensorische terugkoppeling?
2. Vergemakkelijkt een onderdompelende interface het opbouwen van cognitieve kaart?
3. Verstoort discontinue navigatie het bepalen van de huidige locatie?

Om deze vragen te beantwoorden is een serie van negen experimenten uitgevoerd. Hierin wordt de efficiëntie van het bepalen van de huidige locatie onderzocht en wordt gekeken hoe snel een cognitieve kaart wordt opgebouwd, wanneer gebruik wordt gemaakt van de verschillende interfaces en verplaatsingsmethoden.

De resultaten geven aan dat de kwaliteit van padintegratie sterk wordt verbeterd bij gebruik van de onderdompelende interface. Hierdoor wordt ook sneller een cognitieve kaart opgebouwd. Nadat de virtuele omgeving is geleerd worden geen verschillend meer gevonden tussen de onderdompelende en de niet-onderdompelende interface.

Discontinue verplaatsing verstoort het bepalen van de huidige locatie, waardoor moeilijker een cognitieve kaart wordt opgebouwd. Het desoriënterende effect van een discontinue verplaatsing kan deels worden gecompenseerd door voldoende mogelijkheid te bieden om cognitief te anticiperen op de bestemming van de verplaatsing. Het bepalen van de huidige locatie blijft echter het meest efficiënt wanneer continue verplaatsing gebruikt wordt. De tijd die nodig is om cognitief te anticiperen hangt af van het type verplaatsing. Verplaatsingen waarbij een grotere rotatie wordt uitgevoerd en verplaatsingen over langere afstanden vergen meer tijd voor cognitieve anticipatie.

Het ondersteunen van een goede ruimtelijke oriëntatie van een operator is een vereiste stap op weg naar de mogelijke toepassing van VE-technologie als interface voor de supervisie van complexe processen. Bij het kiezen van een technologie zal een afweging gemaakt moeten worden tussen de hier gevonden voordelen van een onderdompelende interface en de bijkomende nadelen. Onderdompelende interfaces zijn aanzienlijk duurder en er zijn diverse bijwerkingen van de technologie die nog niet opgelost zijn. Zo treedt vaak bewegingsziekte op en krijgen veel gebruikers last van hun ogen of krijgen hoofdpijn. Deze beperkingen staan momenteel nog een breed gebruik van de technologie in de weg en moeten eerst opgelost worden.
De resultaten van de experimenten suggereren dat een onderdompelende interface alleen voordeel biedt voor toepassingen waarin herhaaldelijk een nieuwe virtuele omgeving geleerd moet worden, of in toepassingen waar de doelgroep van gebruikers geen ervaring heeft met het gebruik van de technologie. Mogelijke voorbeelden hiervan zijn het leren van de indeling van specifieke gebouwen in een virtuele omgeving door brandweerpersoneel, of het virtueel bekijken van een nog niet bestaand gebouw door potentiële kopers. Voor de toepassing in supervisietaken zal de onderdompelende interface alleen een voordeel bieden wanneer de locatie van informatie nog geleerd moet worden tijdens het bekend raken met de interface. Na deze leerfase mag geen voordeel meer verwacht worden van de onderdompelende interface.
Het discontinu verplaatsen moet niet toegestaan worden tijdens het verkennen van een nieuwe virtuele omgeving, omdat dit het leerproces bemoeilijkt. Deze aanbeveling verandert wanneer de omgeving eenmaal bekend is. Als tijd geen rol speelt dan is continu verplaatsen nog steeds te prefereren, maar als tijd een kritieke factor is dan moet de voorkeur worden gegeven aan discontinue verplaatsing. De desoriënterende werking van een discontinue verplaatsing kan geminimaliseerd worden door cognitieve anticipatie mogelijk te maken. De interface ontwerpen moet zorgen dat voldoende informatie beschikbaar is over de bestemming van de verplaatsing. Discontinue verplaatsingen waarbij een rotatie wordt uitgevoerd moeten zo veel mogelijk worden vermeden.

Verder onderzoek wordt aanbevolen naar de kwaliteit van de ruimtelijke oriëntatie bij meer typen discontinue verplaatsing. Hiermee kunnen meer volledige richtlijnen worden verkregen voor het ontwerpen van interfaces waar discontinue verplaatsingen mogelijk zijn.

Niels Bakker Januari 2001

Curriculum vitae

Niels Bakker was born in Dormaa Ahenkro in the Kumasi region of Ghana, on 26 December 1969. He first received his native name, Kofi Bronja, meaning Friday Christmas, but was later baptised as Nicolaas Hylke Bakker, named after both his grandfathers.

He gained his VWO diploma from the “Barlaeus Gymnasium” in Amsterdam, in 1988. After having studied theoretical physics at the “Universiteit van Amsterdam” for one year he switched to Mechanical Engineering at Delft University of Technology. He obtained his M.Sc. degree in 1995, with a study on the development of a quantitative behavioural model of a tugboat master in the Man-Machine Systems group. After graduating, he worked for one year on new concepts for traffic control of automatic guided vehicles, at the same university in the Logistics group.

In October 1996 he obtained a position as a PhD student for four years within the Man-Machine Systems group, and was posted full-time at TNO Human Factors, in Soesterberg. It was at the TNO Human Factors Virtual Environment lab, that the experiments were performed that are the subject of this thesis. During these research years he co-operated in several projects at TNO, including evaluation studies of the bridge layout of ships using virtual environment technology, and the design and construction of a mechanical head-tracker.

He is currently working as a Post-Doc researcher both withing the Man-Machine System group of Delft University of Technology and at the department of Medical Physics at the Amsterdam Medical Centre. His research focuses on improving the mechanical properties of surgical catheters so that their reach and manoeuvrability can be improved. This research is part of a program on Minimal Invasive Surgery and Intervention Techniques (MISIT) currently continuing at Delft University of Technology.