Individualized Virtual Reality Rehabilitation after Brain Injuries

by

Sebastian Thomas Koenig

A thesis submitted to the faculty of the
University of Canterbury, College of Engineering
in partial fulfillment of the
requirements for the degree of
DOCTOR OF PHILOSOPHY
in Human Interface Technology

Approved by the
Examining Committee:

Andreas Dünser, Senior Thesis Adviser

Christoph Bartneck, Thesis Adviser

John Dalrymple-Alford, Thesis Adviser

Human Interface Technology Lab
Christchurch, New Zealand
August, 2012

ABSTRACT

SEBASTIAN T. KOENIG. Individualized VR Rehabilitation after Brain Injuries

Context-sensitive cognitive rehabilitation aims to address the specific deficits of patients by taking into account the unique strengths and weaknesses of each brain-injured individual. However, this approach requires customized assessments and trainings that are difficult to validate, time-consuming or simply unavailable for daily clinical use. Given the currently struggling economy and an increasing number of patients with brain injuries, a feasible and efficient solution for this individualized rehabilitation concept is needed.

This dissertation addresses the development and evaluation of a VE-based training and assessment for context-sensitive cognitive rehabilitation. The proposed application is designed to closely resemble real-world places that are relevant to each individual neurological patient. Despite such an ecologically valid approach to rehabilitation, the application also integrates traditional process-specific tasks that offer potential for standardization and collection of normative data across patient populations.

Three cognitive tasks (navigation, orientation, spatial memory) have been identified for use in individualized VEs. In three experimental trials the feasibility and validity of the technological implementation and theoretical foundation of these tasks has been assessed. In a fourth trial one of the tasks has been used for the rehabilitation of a brain-injured patient. Based on the results of these studies a workflow for the rapid development of VEs has been established which allows a VR developer to provide clinicians with individualized cognitive tasks. In addition, promising results for the clinical use and validation of the proposed system form the basis for future randomized controlled clinical trials.

In conclusion, this dissertation elaborates how context-sensitive and process-specific rehabilitation approaches each offer a unique perspective on cognitive rehabilitation and how combining both through the means of VR technology may offer new opportunities to further this clinical discipline.

ACKNOWLEDGMENT

Throughout this dissertation I was fortunate enough to be supported by many people, to whom I'd like express my sincere gratefulness. Firstly, I would like to thank my advisors who have given me their fullest support and were always there for me. I am deeply grateful to Dr. Greg Crucian for being an amazing mentor, for giving me all the freedom I needed for my research, and for always believing in me. Most of all, I thank him for being a friend.

Also, I am truly thankful to Dr. Andreas Dünser who has been a great mentor and friend. His door was always open for me and I deeply appreciate the experience, inspiration, and guidance he has shared with me.

I would also like to thank Dr. Christoph Bartneck and Dr. John Dalrymple-Alford for their continuous support and advice. This dissertation would have not been possible without the help and guidance of my advisors. I consider myself very fortunate for having the opportunity to work with such extraordinary researchers.

I would also like to acknowledge the guidance I have received from my former advisors, Prof. Dr. Oliver Tucha and Prof. Dr. Klaus W. Lange, who have strongly supported my studies in Germany and my endeavors to pursue a Ph.D.

Further, I would like to express my gratefulness to Dipl.-Psych. Wolfgang Kühne, who has been a great mentor during my clinical work and the studies conducted at the Asklepios Rehabilitation Clinic in Schaufling, Germany. I would also like to thank all the therapists, clinical teams, and especially the psychology team in Schaufling for keeping me sane during my clinical trials.

I would like to extend my thanks to Professor Dr. Albert "Skip" Rizzo and Professor Dr. Paul Sharkey for their advice, support, and friendship. It is for them that I have found my home in the field of VR rehabilitation.

I would further like to thank my friends, both in New Zealand and Germany, for their patience and support during countless night shifts, data collections, and deadlines. My special thanks are due to Carolin Ritter and Roberto Neubert for listening to my ramblings and putting up with the oddities and stress that these countless hours of work brought to my life.

I would also like to acknowledge the support that the staff and students at the HIT Lab New Zealand and the Psychology Department at the University of Canterbury have given me. I appreciate the thoughtful conversations and brain storming sessions with my fellow students at the HIT Lab NZ and I also thank Dr. Mark Billinghurst for the trust he has given to me and my projects.

I also thank the University of Canterbury, New Zealand, for the generous financial support that I have received through my doctoral scholarship.

Furthermore, I would like to use this opportunity to thank all the participants who have contributed to this dissertation. It is their time-commitment and faith in my research that has made this work possible. I sincerely thank all the patients from Schaufling who have participated in my studies. They have spent

many hours completing my assessments and trainings despite the distressing circumstances that have brought them to a rehabilitation clinic. Their well-being makes all this work worthwhile.

I would also like to express my sincere gratefulness to my fiancée Belinda Lange who has always encouraged and supported me throughout my work. Her advice, inspiration, patience, companionship and love have made this work possible. I value her support, both professionally and personally and I am deeply thankful to be spending my life alongside such a wonderful person.

Lastly, I would like to thank my parents, Thomas and Ute Koenig. Without their support, encouragement, and love this dissertation and any of my work throughout the past years would have been impossible. It is them who have taught me to never give up and try to achieve the best I can in life. I owe them everything and I wish I could find the words to show them how much I love and appreciate them and all they have done for me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

This dissertation was written with readers of medical or engineering backgrounds in mind. Topics of neuropsychology, rehabilitation and software engineering are discussed at various levels of detail throughout this dissertation. The proposed framework is largely beneficial for clinicians and therapists with interest in VR rehabilitation. However, the described approach requires a multidisciplinary team to implement individualized VR rehabilitation in a clinical setting. A software engineer, 3D modeler or technical artist is required to create the individualized scenarios which are used by clinicians to treat and assess cognitive deficits in brain-injured patients.

1.1 Rehabilitation

Brain injuries often have a lasting impact on a person's life, preventing the individual to live independently and engage in activities of daily life. Rehabilitation seeks to ameliorate cognitive and motor functions after brain injury, but even after prolonged, intense training many patients are left with persisting deficits. Brain injuries such as cerebrovascular diseases and traumatic brain injury (TBI) are among the most common causes of death and long-lasting disability in several countries around the world (Johnston, Mendis, & Mathers, 2009). After cancer and heart disease, stroke is the most common cause of death in countries such as the UK, USA, Germany, and New Zealand. Estimated costs for acute and long-term care after brain injuries are among the highest for any health-care related costs in many countries. For example, annual costs for stroke are estimated to be $62.7 billion in the USA, £4.5 billion in the UK, and NZD157 million in New Zealand. Similarly, direct and indirect costs for TBI were valued to be $60 billion in the USA in 2000. Taking demographic models into account, rehabilitation is going to play an even more important role in the future. Population estimates in developed countries predict a sharp decline of the so-called elderly support ratio (Population Reference Bureau, 2010). This ratio quantifies the number of people aged 15-65 years, divided by the number of people aged 65 or older. Consequently, more elderly people are going to seek support through health-care systems around the world. With such substantial impact of brain injuries on the quality of life of millions of people and on economies worldwide, further research into cost-effective and efficient treatments for brain diseases is of high importance.

The heterogeneity of impairments that neurological patients suffer from is a relevant factor for developing and evaluating treatments. Cognitive impairments are a common result of TBI, stroke, cerebral tumors, neurodegenerative diseases (e.g. Parkinson's disease), and many other brain diseases (Vakhnina, Nikitina, Parfenov, & Yakhno, 2009). The severity of each incident largely depends on the size and location of the brain lesion and the individual circumstances of the patient. Age, gender and the existence of risk factors play an important role for the incidence, progression and recurrence of brain diseases. Further, the individual circumstances of each patient are usually taken into account when rehabilitation is planned and carried out. The International Classification of Functioning, Disability and Health (World Health Organization, 2001) combines social and medical aspects for clinical treatment and considers the individual's impairments, participation in society, and contextual and environmental factors. With such a unique perspective on each patient's case, clinical teams can better understand the problems which need to be solved during rehabilitation. Ylvisaker (2003; 2006) describes this approach as context-sensitive rehabilitation and makes a direct comparison to more traditional process-specific cognitive rehabilitation. Process-specificity refers to the cognitive domains of memory, attention and executive function as a basis for cognitive rehabilitation (Sohlberg & Mateer, 2001, p. 7/8). Cognitive deficits in subdomains such as divided attention, short-term memory and executive function are identified during rehabilitation and trained specifically with repetitive use of cognitive tasks. In contrast, context-sensitive rehabilitation takes the impairments of an individual and puts them in the social and functional context of the patient. Cognitive domains are still used to identify impairments, but are then brought into a framework of the individual's participation in society. When Ylvisaker (2003) compares his context-sensitive approach to process-specific rehabilitation, he refers to Sohlberg and Mateer (1989) who describe several distinct cognitive domains and therapeutic approaches for each domain. However, both rehabilitation concepts don't have to be mutually exclusive. In Sohlberg and Mateer's (2001) updated book, the authors put more emphasis on an integrated approach which also includes the individual circumstances of the patient. The authors stress the importance of the patient's integration into a wider community, the priority of function over deficits, and the focus on generalization of rehabilitation tasks, all of which are at the core of Ylvisaker's work. When both concepts are combined, they provide a more comprehensive framework. Cognitive domains are the basis for task-specific, generalizable training which is relevant to the patient's context. Moreover, the patient's individual strengths and weaknesses, the social background and a focus on participation complement the underlying models of cognitive science. Ylvisaker directly compares traditional and context-sensitive rehabilitation and lists assessments and treatments for both. While standard neuropsychological measures are mentioned for both approaches, the author asks for flexibility during assessment and treatment and suggests observation and exploration to take the individual context of the patient into account. However, questions arise on how the standardized nature of many traditional assessments fit into Ylvisaker's concept. His individual approach to rehabilitation demands sophisticated tools that also meet the specific needs of brain-injured individuals. Standard neuropsychological tests and trainings are suitable to test domain-specific cognitive skills, but often may not account for the unique situation of the patient. In order to develop tasks for an individual rehabilitation approach, several aspects regarding their development, validity and clinical use need to be considered.

1.2 Development

In most developed countries health care spending increases at a faster rate than economic growth. Specifically, total health care expenditure as percentage of the gross domestic product has been rising substantially over the last decades (Kaiser Family Foundation, 2011b). Considering this constant rise in health care expenditure (Kaiser Family Foundation, 2011a) in addition to a recent global economic recession, cost-effectiveness has to be of high importance for development and administration of rehabilitative treatments. However, individualized rehabilitation tasks imply a higher cost of creating unique content for each patient. Gathering information about the individual context of the patient requires additional resources. Using such information to provide flexible tasks and training content is even more costly. An individualized set of tasks is not a single development effort, but rather an ongoing adjustment and content generation for each patient. Thus, reusable content that can be adjusted for the individual context of the patient appears to be a cost-effective solution. For example, task-specific trainings can be created for a variety of scenarios, as long as they are flexible enough to guarantee a relevant experience for each patient. In summary, a balance has to be found between operating costs across brain-injured patients and a high degree of flexibility and relevance for each patient.

1.3 Validity

Generally, the validity of a test "[...] refers to the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of the test" (APA, NCME, & AERA, 1999). Specifically, a valid cognitive task provides information about the cognitive functions of a patient. This information is used to draw conclusions about the patient's ability to live independently, return to work, and engage in activities of daily life. However, individualized rehabilitation tasks and traditional neuropsychological tasks seem to differ substantially in respect to different aspects of validity. Several concepts that contribute to the overall validity of a test have been reported in the methodological literature, each separating traditional and context-sensitive tasks.

External validity describes the extent to which task results can be generalized across different settings and situations (Cook & Campbell, 1979). However, the reason for using individualized tasks is not to generalize task results across a wider population or across less meaningful situations. Rather, the clinician attempts to create a unique experience with relevant situations for each individual. Transfer of skills from these unique experiences to a variety of daily-life settings is highly desirable, though it is not always achievable (Ylvisaker, 2003).

Several clinical trials attempted to demonstrate such skill transfer with traditional rehabilitation methods, but results have been mixed. For example, Ben-Yishay and colleagues (1987) found specific training effects for a training of 40 brain-injured patients using their Orientation Remedial Module (ORM). In a sub-sample of eleven patients, only specifically trained attention domains improved while other untrained domains did not improve. Moreover, the authors were able to show (mostly weak) relationships between the patients' performance on the ORM and activities of daily life.

Cicerone and colleagues (2000) conclude in their literature review that cognitive rehabilitation should generally be directed towards improving everyday functioning. However, only few of the reviewed studies showed any transfer towards the daily life of patients. Some evidence was found for visuospatial rehabilitation, language and communication training and training of compensatory memory strategies.

In an updated review, Cicerone and colleagues (2005) come to similar conclusions and summarize that strategy training generally appears superior to targeting specific skills, especially for memory and attention training. The authors note that future research and clinical practice should pay close attention to functional outcomes and the participation of patients in their social context, rather than train and evaluate patient performance at the impairment level. As in their previous review, evidence for skill transfer to daily activities is scarce.

Geusgens et al. (2007) reviewed 41 studies specifically looking for transfer effects during cognitive rehabilitation. They only included studies that trained compensation strategies as opposed to cognitive skills training. They refer to existing studies that strategy training (i.e. compensating behaviors) is more likely to evoke transfer to the patients' everyday life than directly training the deficient cognitive domain (Cicerone, et al., 2005; Wilson, 2000). Out of the 41 reviewed studies, 36 were able to demonstrate some form of transfer. However, only 22 studies actually evaluated transfer to daily-life activities while the others looked at either simulated lab-based activities or activities that were very similar to the previously trained ones. Out of these 22 studies, 18 were able to show transfer of learned abilities, but only six included statistical evidence for their results. Furthermore, the sample sizes of most studies were very small or based on single-case designs. Consequently, no clear-cut conclusions for or against strategy training transfer to daily activities can be drawn.

In summary, generalization of skills during cognitive rehabilitation towards daily-life settings has only received little support in the literature. More specifically, task-focused training appears to show no transfer to situations outside of the training situation and the effectiveness of strategy training requires further evidence. While the external validity of training applications seems to be of central importance to the patients' success in their daily life, most traditional rehabilitation studies have not successfully demonstrated such transfer yet. Even though principles of context-sensitive rehabilitation have been mentioned in several literature reviews (Cicerone, et al., 2005), context-sensitivity is often not associated with transfer to activities of daily life. This is because context-sensitive tasks are essentially based on the unique experiences that a patient has in his daily life. Hence, a transfer is often not necessary as training tasks are either identical to common daily chores or replicate them as closely as possible. Nonetheless, when traditional process-specific tasks are combined with individualized context, task generalization across similar daily activities seems to be of relevance.

Internal validity is concerned about the causal inferences that can be drawn from task results (Cook & Campbell, 1979). A highly controlled and standardized testing situation may yield results of high internal validity, but has little in common with the everyday situations that a patient is faced with. Again, individualized tasks are aiming to provide relevant situations that might not always be fully controlled by the clinician. For example, a patient could be sent to the hospital's cafeteria to purchase specific items. During this task, the patient might encounter different people, distractions and obstacles each time the task is administered. Hence, internal validity of such task results may vary widely across individuals and trials.

Campbell and Fiske (1959) discussed convergent and discriminant validity as an additional test validation method. In more recent publications, both validity concepts have been summarized as evidence based on relation to other variables (APA, et al., 1999). Task results are correlated with other well-validated tests that rely on similar or different underlying concepts. High correlations are expected for similar tests, whereas tests with different conceptualization are expected to show lower or no correlations. Using unique tasks for each individual patient does not provide a basis to compare to already-established tests, so that convergent and discriminant validity may not be suitable to validate the concept of a context-sensitive task.

Nadolne and Stringer (2001) evaluate the importance of ecological validity for clinical assessments. The authors argue that traditional paper-and-pencil measures do not relate strongly to real world tasks. In their sample of 31 stroke patients, ecological simulations of orientation behavior showed higher correlations with real world wayfinding than traditional tests of visualization and orientation. The concept of ecological validity, which expresses how closely test and real world situations are alike, is arguably a key factor for using individualized rehabilitation tasks. Such tasks' unique and relevant content is aiming to be very close to the patient's everyday problems.

Development of cognitive tests and trainings commonly demands a comprehensive evaluation of validity, including evidence across several different domains of validity (APA, et al., 1999). Only ecological validity, the major advantage of individualized tasks, is not essential for the overall validity of an evaluation study (Shadish, Cook, & Campbell, 2001). In summary, individualized tasks do not seem to be suitable for evaluating all aspects of validity such as external and internal validity. Ylvisaker (2003) also addresses the controversy of scientifically evaluating the flexible treatments of context-sensitive rehabilitation. He admits that a rigorously controlled clinical trial is difficult to achieve due to the uniqueness of each patient's intervention. However, Ylvisaker notes:

> *[…] if context is considered the independent variable, with one group receiving cognitive retraining delivered by rehabilitation specialists in a clinical setting using training tasks that are not individualised and the other group receiving services delivered by everyday people (e.g., family members, teachers, job coaches) in everyday settings with training and support from specialists — and with tolerance for considerable variation in the specifics of that intervention — then the study is conceptually simple. (pp. 11/12)*

The question arises whether this broad evaluation between context-sensitive treatments and traditional alternatives is suitable to identify specific factors that make either approach successful. As previously discussed, the wide range of uncontrolled factors in context-sensitive treatments results in very low internal validity, so that conclusions about its efficacy cannot be attributed to any particular aspect of the treatment other than its individualism for each patient. Moreover, if context-sensitive and process-specific aspects are to be combined in a cognitive task, construct validity and evidence based on internal structure (APA, et al., 1999) play an important role for task development and evaluation. Process-specificity requires that the developed tasks actually measure the underlying cognitive processes that they purport to measure. Consequently, task development should emphasize standardization for strict validity evaluation while still being flexible enough to meet the needs of each brain-injured individual. This controversy emerges as one of the main challenges of this dissertation.

1.4 Clinical Usage

Whenever new treatments are developed it is important to consider the context in which they are going to be applied. Thus, a rehabilitation task has to satisfy a range of requirements that are inherent to a clinical setting. Time is a sparse resource during rehabilitation. On average, stroke patients spend 15 days in the USA (Conroy, DeJong, & Horn, 2009) and three to four weeks in Germany (betanet, 2011) in inpatient rehabilitation. Individualized trainings need to effectively use this time without requiring too many supervised therapy sessions. Due to the complex nature of brain injuries, patients with stroke or traumatic brain injury often require a combination of physio-, occupational and cognitive therapies (Mercier, Therese, Hebert, Rochette, & Dubois, 2001). Hence, a cognitive task should be flexible enough to fit any patient's therapy schedule. As a patient recovers throughout the rehabilitation process, therapy goals and demands change. Any treatment should be able to accommodate these changes and adapt in frequency and difficulty. The short duration of inpatient rehabilitation and limited time of therapists also give rise to the use of unsupervised training that can be continued after therapy sessions have finished or the patient is discharged from inpatient services. In addition, tasks need to be as intuitive to use as possible so that therapists and patients do not require lengthy instructions and supervision. Moreover, motor and cognitive deficits of neurological patients need to be taken into account for usability and accessibility evaluation.

Context-sensitive rehabilitation promotes the use of meaningful, individualized tasks in inpatient and outpatient settings (Ylvisaker, 2003). However, frequent visits of rehabilitation experts to the patient's home or workplace are costly and time-consuming. Additionally, institutionalization after neurological deficits is a common situation in which patients are confronted with an unfamiliar environment (New Zealand Guidelines Group, 2006). These patients are often faced with more severe cognitive deficits (Patel, Coshall, Rudd, & Wolfe, 2002) that prevent them from leaving the hospital and visiting relevant environments during inpatient rehabilitation. Thus, it is important to integrate information of relevant environments into the patients' therapy schedule while they are still at the hospital. However, this process appears unrealistic in many clinical settings to date, as the additional effort to gather the necessary information has to be seen in the context of decreasing health-care budgets. Both, the cost-effectiveness of the treatment and the inclusion of individual information, have to be taken into account when designing a context-sensitive rehabilitation task.

Further, rehabilitation hospitals are often well-structured environments in which patients participate in carefully-designed therapy programs. Therapy frequencies, scope and variety of activities at rehabilitation hospitals are often based on strict standards of health-care providers and rehabilitation guidelines (e.g. International Classification of Functioning, Disability and Health). Therapists and nurses are available around the clock to assist patients with their problems. Such structured settings may not always provide the means to practice relevant everyday tasks under realistic, unstructured and unpredictable conditions. As a consequence, hospital-based individualized tasks need to replicate, or at least approximate, the complexity of everyday situations that patients are faced with outside of the rehabilitation hospital.

Lastly, rehabilitation needs to address the individual's reintegration into the community and social context. Involving family, community service agencies, and schools has been recommended by Ylvisaker (2003), Sohlberg and Mateer (2001), Sloan and colleagues (2004), and Wade (2000, 2001). These parties are directly involved in the care and social reintegration of the brain-injured individual long after inpatient and outpatient rehabilitation have finished. Their support is of critical importance for the development of individualized cognitive tasks, as they can provide much information about the problems and situations that need to be trained during rehabilitation.

In conclusion, context-sensitive rehabilitation has been advocated by several clinical researchers in the past (Adams, 2003; Ylvisaker, 2003). Its implementation in the traditional context of process-specific rehabilitation (Sohlberg & Mateer, 2001) especially in terms of treatment validation and cost-effectiveness must be subject to further investigation. In particular, this dissertation addresses the development of flexible clinical treatments and how they meet the demands of context-sensitive and process-specific rehabilitation.

1.5 Virtual Reality (VR)

Thomas Furness III. (1992) defines VR as “the representation of a computer model or database in the form of a system of virtual images which creates an interactive 3D environment which can be experienced and/or manipulated by the user.” (p. 12).

This thesis proposes a framework of context-sensitive applications which make use of VR technology. Cognitive tasks are embedded in virtual environments (VEs) which can be manipulated by a brain-injured patient. Hence, an introduction to VR is given in the following chapter. While not all of the mentioned applications strictly involve VR technology with a high degree of interactivity and immersion, all of them offer viable approaches to treating cognitive deficits and provide valuable information for the clinical application proposed in this dissertation.

Cognitive training applications such as Rehacom[1], CogniPlus[2] and CogPack[3] provide a wide range of tasks for process-specific rehabilitation. They promote high-frequency training for specific cognitive domains. These cognitive domains are based on scientific models, such as van Zomeren and Brouwer’s (1994) model of human attention. A large selection of individual tasks is aimed at training specific cognitive sub-systems as described in the underlying model (e.g. divided attention, sustained attention). The therapist chooses which tasks are the most appropriate for the patient’s individual situation and deficits. This approach allows for individualized treatment for each patient through selection of relevant tasks. While some tasks have been designed to resemble common situations and chores that people are confronted with in their daily life (e.g. driving a car), others are variations of puzzles and mini games (e.g. card games) that do not have to be relevant to the patient’s context. Furthermore, only few clinical trials have shown their effectiveness beyond traditional therapy concepts. Most literature overviews on the publishers’ websites list studies in psychiatric rehabilitation of which most have been conducted more than a decade ago (Olbrich, 1996). However, technology has changed substantially since then and several of the listed programs have long been revised or support of previous versions has been discontinued.

RehaCom also presents a list of clinical trials in which their training has been used. Exemplarily, Weiand (2006) examines the efficacy of a process-specific training based on RehaCom tasks and compares it to an unspecific training of similar complexity using Microsoft Word. 51 patients were included in this study and randomly given either an unspecific training using Microsoft Word or a specific attention training using the RehaCom software. A first supervised training consisted of three weeks and a total of 15 hours of training. A second training

---

[1] Hasomed RehaCom – www.hasomed.de

[2] Schuhfried GmbH – CogniPlus – www.schuhfried.at

[3] Marker software – Cogpack – www.cogpack.de

protocol allowed the participants to continue their previous training method on their own for additional three weeks. Neuropsychological assessment and an assessment of activities of daily living were conducted before and after the training protocols. The effect of both trainings was examined regarding to specificity, transfer, generalizability and motivation to continue training voluntarily.

A total of four cognitive tasks of the RehaCom software were used during the specific training protocol. The tasks “AUFM” (attention and concentration) asks the participant to match series of pictures to a reference picture. The reaction time task “REAK” was used to train reaction to visual and acoustic stimuli in an individual or forced choice paradigm. The “VRO1” tasks trains spatial abilities by displaying two-dimensional stimuli which have to be matched to a reference image. The displayed images contain features which require aspects of mental rotation. Lastly, the “GEAU” task places demands on the user’s divided attention by displaying several sets of visual cues to which the user has to react as quickly as possible.

Despite the lack of adjustment for multiple statistical comparisons (i.e. Bonferroni adjustments), the author does not find any significant differences between both training approaches, as both groups show large improvements across most attentional outcome measures. Further, no tangible transfer of the training to the patients’ participation in daily-life was found. Any improvements on the used self-report measures of activity and participation in life can more plausibly be explained by conventional physio- and occupational therapy that the patients received. However, the process-specific RehaCom tasks appear to be successful at keeping patients motivated for continuous training even after the supervised sessions at the clinic have finished. As such, the process-specific training seems to be a good choice for long-term self-guided exercises. It just remains to be tested whether a more individualized training approach can actually be superior to the tasks that Weiand (2006) used. The lack of meaningful evidence for the training's transfer to daily situations further underlines the need to evaluate context-relevant training tasks.

It has to be noted that not all of the available programs provide normative data for their respective subtasks (e.g. RehaCom), as they are primarily designed for therapeutic use, not for cognitive assessment. On the contrary, computerized neuropsychological assessments require a thorough evaluation of validity and reliability. For example, Testbatterie zur Aufmerksamkeitsprüfung[4] and Wiener Testsystem[5] assess distinct cognitive functions by using a battery of abstract tasks. Both systems provide normative data for several age groups. Numerous clinical trials have taken advantage of the process-specific nature of either test by assessing subdomains of attention (Sturm, Willmes, Orgass, & Hartje, 1997; Tucha et al., 2008) or cognitive performance for driving assessment (Golz, Huchler, Jörg, & Küst, 2004). Again, only few of the provided tasks resemble everyday situations that are meaningful for the unique background of patients with brain injuries. Patients

---

[4] PsyTest – Psychologische Testsysteme – www.psytest.net

[5] Schuhfried GmbH – Wiener Testsystem – www.schuhfried.at

mostly have to react to abstract stimuli that are presented on screen. For example, TAP's divided attention task displays an array of crosses on the screen and the user has to push a button whenever four crosses form a square. As a second task low and high tones are presented to which the user has to respond whenever two of the same tones are presented consecutively. While this task is based on existing theories of divided attention (Zomeren & Brouwer, 1994), it does not closely resemble any meaningful tasks that a patient could be confronted with in a home environment. However, as these computer tasks are only used for cognitive assessment, in some instances fidelity and ecological validity can be of less importance than for a comprehensive training program. In fact, several computerized attention assessments purposely choose abstract tasks or place specific demands on the attentional system. For example, TAP's vigilance test simulates a very monotonous task by repeatedly alternating a visual stimulus between two locations on screen. In very few occasions the user has to press a button whenever the stimulus remains on the same location for two consecutive trials. As long as the validity and reliability of such tasks has been based on results of methodically sound research studies, such abstract tasks can provide a basis for the clinician's decision-making process throughout cognitive rehabilitation. Taken together, computerized assessment and training tasks are valuable tools as long as they are well-grounded in theories of cognition and possess excellent psychometric properties. However, due to their low fidelity and ecological validity they may not be an optimal choice to assess and train patients during cognitive rehabilitation when the patient's unique context is a major concern.

Looking at context-sensitive rehabilitation, very few computer-based tasks have utilized individualized content. NeuroVR (Riva et al., 2011) has been among the first to actively pursue individualized training and therapy content. When NeuroVR 1.0 was first tested to evaluate its feasibility for individualized cognitive rehabilitation in 2008, the application was very limited in its fidelity, customizability and user interface. Since then, Version 2.0 has been released in early 2011 with a wider range of integrated environments and support for importing external 2D and 3D models. Generally, the program is designed to enable therapists to create meaningful VEs for clinical applications. For example, the authors of the software have used NeuroVR to treat obesity and anxiety disorders (Gorini & Riva, 2008). The scenarios usually consist of existing scenes and imported media (e.g. sounds, pictures, 3D models). However, NeuroVR is not an out-of-the-box therapy tool; it rather provides a toolset to create meaningful experiences for patients. Several generic VEs are included in the software and can be used and modified as needed for a therapy session. Despite its recent upgrade to Version 2.0 the program's visual quality is still low compared to modern game engines. For example, the rendering engine does not support advanced features such as real-time shadows, ambient occlusion or postprocessing effects when compared different applications described in later chapters. Also, the setup of a meaningful environment requires substantial work. Basic interactivity is provided within the scene editor. While this set of 'triggers' can replicate straight-forward tasks of picking up items, playing sounds and animations, more sophisticated cognitive tasks will require access to more functionality or the editor's source code. Essentially, NeuroVR can be seen as a simulation engine with limited access to features such as rendering, file import/export, user interface and scripting. The scripting of interactivity has been implemented as an icon-based interface. However, without knowledge of scene editors and their functionality substantial time-investment can be required to build a relevant virtual scenario.

Taken together, NeuroVR is a tool to create, edit and present VEs for therapeutic use. Compared to more complex game engines, it provides a more accessible entry for using VEs in therapy. However, each editor program needs to make a tradeoff between ease of use and complex functionality. NeuroVR in its current version (Version 2) appears to be primarily developed for clinicians who seek to use basic VEs for their treatment. Thus, the flexibility and complexity of the produced scenarios is limited and may not always be well-suited for complex neuropsychological assessments and trainings. This aspect is critical for the implementation of a set of well-defined cognitive tasks in VEs. Without the flexibility to precisely control displayed content, user interface and data recording, an iterative patient-centered development process does not seem feasible with NeuroVR. Such flexibility and the intended clinical use are the main aspects that set NeuroVR and the application developed during this dissertation apart. NeuroVR is presented as a development tool for clinicians. On the contrary, the proposed application of this dissertation contains cognitive tasks which are embedded in individualized VEs. Task development and implementation of relevant task context are done by a software developer using a modern game engine to increase the application’s flexibility and fidelity. The resulting tasks are developed iteratively based on feedback of patients and clinicians. The finished individualized tasks are then used by a clinician who is not required to invest additional time into development or integration of individualized context. Most importantly, each VE is identical to the relevant real-world scenario instead of being a generic scene enhanced through familiar sounds and objects.

Several other computer-based programs for treatment of anxiety disorders, phobias and post-traumatic stress disorder (PTSD) have been developed in the past. All of these applications are to some degree flexible and adjustable to the patient's needs. While therapy scenarios generally are static, the therapist can control the patient's exposure to critical stimuli. For example, VEs for exposure therapy have been successfully used for combat-related PTSD (Rizzo et al., 2010), PTSD caused by terrorist attacks (Freedman et al., 2010), spider phobia (Garcia-Palacios, Hoffman, Carlin, Furness III, & Botella, 2002), Aviophobia (Price, Anderson, & Rothbaum, 2008) and Agoraphobia (Vincelli et al., 2003). The therapist's influence on stimulus exposure can vary considerably between applications. Virtual Iraq (Rizzo, Graap, et al., 2009) and Virtual Afghanistan (Rizzo, et al., 2010) provide for a high degree of customization to allow the therapist to re-create the patient's traumatic experiences for controlled exposure. Complex visual stimuli, sounds and odors can be produced on demand in response to the patient's vebal feedback or physiological parameters such as heart rate and galvanic skin response. This individualized, immersive therapy is able to achieve positive outcomes that go beyond the results of traditional imagination therapy (Rizzo, Difede, et al., 2009). Exemplarily, Rizzo and colleagues (2009) found significant and clinically meaningful decreases in scores of the PTSD Checklist-Military Version and the Beck Anxiety Inventory such that 16 of the 20 participants did not meet diagnostic criteria for PTSD after the VR treatment protocol.

Customized VEs have also been used in psychological experiments, especially in the domain of navigation research. Real-world places have been modeled and turned into virtual research scenarios for participants to walk through. Even though these environments are often not flexible and customizable as such, they have been specifically built to resemble real environments. Koh and colleagues (1999), Ruddle, Payne and Jones (1997) and Witmer, Bailey, Knerr and Parsons (1996) successfully applied VEs for studying human navigation and were able to demonstrate the utility of such environments for training purposes. The authors' results suggest that the simulation of real-world scenarios is feasible so that similar technology could also be used to create realistic, individualized VEs for cognitive rehabilitation.

When considering VEs for context-sensitive rehabilitation, it is important to first evaluate the limitations and potential of the underlying VR technology. For the past decades VR technology has been used in many different domains such as education (Virvou & Katsionis, 2008), simulation for expert training (Lewis, Aggarwal, Rajaretnam, Grantcharov, & Darzi, 2011) and therapy. Looking at medical uses in particular, Rizzo and Kim (2005) and Rizzo, Schultheis, Kerns and Mateer (2004) discuss the advantages and disadvantages of VR systems in a therapeutic context. Even though both reviews have been conducted six and seven years ago respectively, most of what the authors discuss still appears to be of relevance. In Rizzo and Kim’s overview the following aspects were among the key characteristics for VR systems and therefore should be taken into account when developing VEs for individualized rehabilitation.

1.5.1 Systematic and controlled delivery of complex stimuli

One of the biggest advantages of VR applications is the possibility to create large and complex environments while still being under control of every aspect of the system. Stimuli can be presented systematically and timed precisely. This feature is of high importance for development and evaluation of rehabilitation assessments where repeatability and standardization are critical factors. The resulting applications can be of high internal validity and reliability without compromising the complexity of the delivered stimuli. For example, Rizzo and colleagues (2010) developed a complex scenario for treatment of combat-related post-traumatic stress disorder. Several simulated scenarios are embedded in realistic Iraqi or Afghanistan environments. Despite the environment’s complexity, the therapist is under precise control of the exposure and severity of the presented stimuli so that patients can be gradually confronted with stress-inducing situations.

The standardization of tasks within individualized environments also provides the opportunity to compare patient performance with normative data. Further, task performance can be compared across patients and training sessions to quantify training progress. By keeping most task parameters constant and manipulating specific stimuli, the therapist can selectively test scenarios depending on the patient’s individual needs. This methodology essentially allows for hypothesis-driven evaluations for individual brain-injured patients or even comprehensive clinical trials.

1.5.2 Enhanced ecological validity

As previously outlined, ecological validity can be seen as a key component for assessing cognitive skills that are relevant for functional tasks in a real-world context (Nadolne & Stringer, 2001). For the purpose of this discussion enhanced ecological validity relates to a comparison to traditional paper and pencil assessments or their computerized counterparts. Task transparency and relevant functional tasks such as wayfinding through a VE or remembering groceries for preparing a breakfast in a virtual kitchen, are examples where ecological validity can be described as enhanced when compared with abstract traditional assessments of cognitive functions.

A constant rise in processing power enables modern computers to render VEs very realistically. Accurate physics simulations, realistic lighting and human-like avatars provide experiences that are close to the real world. Immersive displays and intuitive interaction methods further enhance the user's experience so that ecologically valid scenarios can be created that closely resemble relevant, naturalistic settings. With growing popularity of video games (Entertainment Software Association, 2007) more tools and resources are becoming available to develop VEs with even higher levels of realism. Modern game engines already provide the technology to develop environments that can be easily recognized by users and allow for high visual quality. Trenholme and Smith (2008) give an overview of several game engines and their functionality for the development of first-person VEs. Since then, several other game engines such as UDK[6], Torque3D[7] and Unity[8] have become available to produce interactive 3D environments of even higher quality.

Already more than a decade ago VEs provided fairly realistic simulations in which users could learn spatial layouts and apply their knowledge to real-world places (Ruddle, et al., 1997). More recently, realistic applications have been developed for treatment of post-traumatic stress disorder. The technology behind Virtual Iraq (Rizzo, Graap, et al., 2009) and Virtual Afghanistan (Rizzo, et al., 2010) is based on the video game “Full Spectrum Warrior” by Pandemic Studios. The VEs' ecological validity is enhanced through the use of head-mounted display, haptic feedback, realistic sounds and exposure to odors. The different modalities are expected to trigger memories and consequently a stress reaction in the user in order to successfully apply methods of cognitive behavioral therapy.

For the purpose of this thesis it is assumed that the visual quality and realism of the VEs are of central importance in order for patients to recognize and acknowledge the relevance of the task and context at hand. Essential characteristics of virtual scenarios and tasks (i.e. transparency, believability,

---

[6] Epic Games – Unreal Development Kit – http://udk.com

[7] GarageGames – Torque3D – www.garagegames.com

[8] Unity Game Engine – www.unity3d.com

plausibility, and relevance) are summarized under the term “realism” in order to describe that the patient can recognize the employed tasks and scenarios and refer to them based on past experiences. Using real photographs for texturing and remodeling furniture and accessories strongly enhance the realism of the environment. The relevant virtual scenario captures the patient’s interest and improves long-term motivation to use the virtual tasks at high frequencies.

Transparency and “realism” in a broader sense can relate to plausibility and place illusions which are described by Slater (2009). Plausibility illusion refers to the fact that the user believes the virtual scenario is actually occurring. It is caused by events and the scenario relating directly to the user (e.g. virtual character talking to user). Place illusion refers to the sensation that the user is actually situated in the displayed location and is described in relation to sensorimotor contingencies of the VR system (e.g. user interaction, tracking and multimodal user feedback). Slater’s definitions do not exactly fit the scenarios of this thesis as there are no virtual events directly targeted at the user and simple desktop systems are being used. However, task transparency and relevant virtual scenarios are believed to contribute to the described illusions that virtual events and locations are actually relevant for the user and engaging for cognitive rehabilitation. For example, a cognitive task that is embedded in a user-relevant scenario directly relates to the therapy goal of the patient and represents a desired outcome of the patient’s rehabilitation (e.g. virtual kitchen with cooking tasks relates to the scenario that the patient aims to engage in independently at home). This stands in contrast to the abstract nature of traditional neuropsychological tests which may have little in common with real-world scenarios (e.g. using abstract objects for mental rotation). Scenarios of high realism are believed to be of advantage when patients deny their cognitive deficits. The realism of a task can potentially lead patients to compare their performance with common standards and past experiences and make them realize that their cognitive abilities may not match their subjective perception. This is the basis for patients actively engaging in cognitive training and making progress throughout their cognitive rehabilitation.

1.5.3 Immediate performance feedback

VEs are capable of delivering automated feedback depending on the user's responses. Feedback can be provided about the quality of the patient's performance whenever tasks and problems are solved. More importantly, dynamic feedback can guide the user during a task to promote error-free learning. For example, sounds or visual cues can be triggered based on the user's movement, distance towards targets, or any arbitrarily defined parameter within a VE.

1.5.4 Ability to pause and resume assessments

Closely related to feedback delivery is the ability to pause and resume assessments and trainings at any time. In real-life scenarios or traditional neuropsychological tests it is often not possible to leave the current testing situation or to interrupt a testing session at any time. VR applications can often be paused and resumed as needed in order to explain strategies and give verbal feedback to the patient. Immediate breaks during training sessions might also be necessary during exposure therapy. The virtual scenario can be stopped quickly when the patient's stress-level increases.

1.5.5 Extensive capabilities for recording and analyzing user behavior

Sophisticated collection of user data is a major advantage of VEs. Modern VEs are often based on realistic physics-models and built to scale. Every object and avatar is placed in a three-dimensional coordinate system which allows for precise measurements of all user movements within the VE. These measurements can then be used to compare against or predict real-world performance. Further, the user's interactions with the application via a user interface can be recorded and analyzed. Eye gaze, body tracking, electroencephalography (EEG, i.e. brain-computer interface) and psychophysiological measures such as heart-rate and galvanic skin response can provide information about the user's cognitive and emotional processes. Collected data can often be processed in real-time and used for direct feedback or graphical analysis. In sum, there is a multitude of data collection tools and methods available that can be employed within VEs or through means of external hardware. Such data collection abilities are not exclusive to VR applications, but can be seen as an advantage when used in conjunction with a VE. Specifically, stimulus exposure can be controlled precisely in virtual scenarios so that the occurrence of a critical stimulus can be recorded and related to the user's reaction (e.g. EEG, psychophysiological measures). This procedure can provide additional insights into neural, physiological and behavioral aspects of human performance when compared to traditional paper and pencil assessments during which stimulus exposure cannot be timed and recorded as precisely.

1.5.6 Backend data extraction and management

The large amount of data that can be extracted from a VR application can exceed the complexity of results of traditional neuropsychological assessments. For example, extensive log files of timed events within a VE are contrasted with simple reaction times or correctly/incorrectly answered items on traditional paper and pencil assessments. It then becomes a question of how this complex data can be processed, stored and condensed to aid clinicians in their decision-making process. Clinicians often do not possess advanced knowledge in software engineering to understand the underlying design, development and capabilities of the used application. Hence, interactions with VR applications should be as intuitive as possible without requiring much technical knowledge or programming. It is also advisable to integrate capabilities of data analysis and visualization into the developed VR application in order to avoid the need for additional software that the clinical user needs to handle. Such analysis should provide comparisons to normative data and a variety of scores that are usable in a clinical context (Rizzo & Kim, 2005). Taken together, a balance needs to be found between the complex data that can be output by modern VR applications and a user-friendly overview of results that takes into account the clinician's needs. Large data sets can be useful to extract information which has not been available in traditional assessments. However, a clinician often may need to make a decision about whether a patient can live independently (or return to work, return to duty) and hence should only be exposed to as little data as is needed to make a well-informed decision.

1.5.7 Delivery of safe and risk-free training environments

VEs can be used to simulate a wide range of scenarios. Training environments can either be inaccessible for real-world training or simply too dangerous to practice in. For example, VEs can be used to safely expose at-risk-populations to critical tasks like street-crossing behavior (Katz et al., 2005). Surgeons or medical students can practice complicated procedures before applying their skills on real patients (Parsons et al., 2008). Patients during cognitive rehabilitation often have no access to relevant training environments while they are at a rehabilitation hospital. After suffering from stroke or traumatic brain injury, patients may spend several weeks in acute and rehabilitation clinics without returning to their home or workplace. Home visits are often not allowed due to safety concerns. VEs can then provide safe and meaningful training content to these patients.

1.5.8 Accessibility for users with motor and sensory disabilities

Many brain-injured patients suffer from motor or sensory impairments (Walker & Treven, 2007) that prevent them from using traditional input devices such as mouse and keyboard. During assessment and therapy these impairments can prevent patients from receiving beneficial treatments or confound results of traditional tests. VR applications have attempted to adapt to the impaired user base by implementing novel interaction methods. Joysticks, modified gamepads, eye-tracking or speech-interfaces have been used in various trials and rehabilitation programs. More recently, brain-computer-interfaces, gaming input devices and body tracking have become increasingly popular and affordable to make rehabilitation tasks accessible for patients with disabilities. The Nintendo WiiMote has been particularly popular with researchers and clinicians as it provides a cost-effective, intuitive input device for patients of all ages (Lange et al., 2010). An even more intuitive solution provides the recently published Microsoft Kinect[9] for full-body tracking (Lange, Rizzo, Chang, Suma, & Bolas, 2011). With its tracking capabilities it even enables patients with severe motor disabilities to move through VEs who otherwise could not participate in most rehabilitative treatments.

1.5.9 Motivating nature with gaming content

Realistic VEs and elements of gameplay can increases motivation for continuous training over extended periods of time (Prensky, 2002). Introducing characters, adding achievements, scores and telling a story can be used to distract the patients from the fact that they are being tested (Rizzo & Kim, 2005). The importance of training motivation becomes apparent when looking at rehabilitation research. High-frequent repetition of rehabilitative tasks has been suggested to promote amelioration of cognitive and motor functions after brain damage. Moreover, task repetition in different contexts appears to be critical to promote generalization of practiced skills (Sohlberg & Mateer, 2001, p. 20). However, frequently repeating monotonous tasks or going through abstract assessment batteries can adversely

---

[9] Microsoft Kinect – www.kinectforwindows.org / www.xbox.com/kinect

impact a patient’s motivation. As the growing number of serious games suggests, engaging game-like training content appears to be a method of choice to prevent frustration and boredom of users.

1.5.10 Lack of tools and standards for development process

Development of therapeutic VR applications requires knowledge in software engineering, VR hardware (e.g. Head-Mounted Display, CAVE displays, tracking systems), usability and rehabilitation/clinical sciences. Each of these domains comes with its own challenges. For example, there is an abundance of software development suites available that can be used to create VR applications. Some of them were not specifically designed to develop VR applications (i.e. game engines such as Unity3D, Torque, Ogre3D, UDK) and others already provide support for VR hardware (e.g. Quest3D, Virtools, NeuroVR). In addition to these “engines”, 3D modeling programs and middleware applications (e.g. for physics or networking) are sometimes necessary to implement features that are needed for individual projects. Integrating all these engines and tools to develop rehabilitation software often results in “one-off” projects that are too complex to adapt them to different user groups or re-use them in several hospitals with larger groups of patients. Consequently, a goal of clinical VR system development should be a transparent application which can easily be adjusted to the hardware configuration of different hospitals and the needs of different patient groups (Rizzo & Kim, 2005).

In addition, the process of user-centered design needs to take into account both user groups of clinical VR applications: patients and therapists. Great care needs to be taken when VR hardware and input devices are implemented for neurological patients. There is a large heterogeneity within and between different patient populations (e.g. frontal-lobe damage vs. right-hemispheric stroke) regarding their motor, sensory, and cognitive deficits. Thus, a wide variety of input devices and displays needs to be considered and tested to avoid ethical (e.g. stereoscopic displays for epilepsy patients) and usability issues (e.g. tremor patient using mouse input).

1.5.11 Adverse side effects

Side effects such as simulator sickness are still a problem for the wide-spread use of VR applications. Even though with powerful hardware performance lag is becoming less of an issue, some participants are still affected by symptoms of simulator sickness when using unintuitive and complex interfaces (Stanney, Kingdon, Graeber, & Kennedy, 2002). To minimize the risk for patients Stanney's (2002, pp. 721-730) guidelines for exposure to VEs are still the most accurate and up-to-date protocols available for development of VR applications.

1.6 Summary

In summary, VR applications possess a large potential for systematically delivering realistic training scenarios to patients undergoing neuropsychological rehabilitation. Primarily, it is the complexity of the development process that appears to be a limiting factor for using VR applications for individualized rehabilitation. Ongoing costs for modeling unique 3D content and programming individualized tasks need to be kept low for widespread use in clinical settings. Moreover, if the applications' side effects can be minimized and the development process adjusted to the short time-frame of inpatient rehabilitation, VR technology could provide a powerful alternative for individualized cognitive rehabilitation.

1.7 Research Overview

It is the aim of this dissertation to develop, validate and apply a framework for cognitive tasks that is based on the strengths of VR-technology and meets the demands of a clinical rehabilitation context. Specifically, several standardized cognitive tasks have been developed that are embedded in unique VEs. Three tasks have been identified and developed within the scope of this dissertation. While the complete rehabilitation system is expected to encompass several tasks from each cognitive domain (memory, attention, executive functions), this dissertation is aiming to evaluate the task domains of navigation, spatial orientation and spatial memory. Tasks were chosen for their relevance in neurological rehabilitation – specifically for patient with stroke for whom it is essential to be spatially oriented in order to live independently.

Each VE is individually created to represent the user's relevant context. Therefore, it is expected that the environment is meaningful to the brain-injured patient. This approach draws upon elements of process-specific and context-sensitive cognitive rehabilitation and provides a connection between both through the means of VEs. Thereby, the rapid development process and the standardization of the cognitive tasks constitute major improvements upon existing VE-based rehabilitation programs. The modularity of the cognitive tasks reflects the standards of current process-specific assessments (Sohlberg & Mateer, 2001). The individual environment in which the tasks are embedded adds personal context to the task. This procedure is based on practices of context-sensitive rehabilitation suggested by Ylvisaker (2003).

For the purpose of evaluating the three cognitive tasks, four experiments are described throughout this dissertation. Each experiment uses realistic VEs which are based on real-world places. However, only Experiment four features an individualized environment which was specifically created for the user. During the early stage of this dissertation, Experiment one was conducted to test the Virtual Navigation Task (VNT) in a real and VE. In order to implement this task in a rehabilitation context it needs to be shown that navigation in virtual and real environments is similar. Consequently, a virtual navigation task could be meaningful for clinical decision-making about real-world behavior. Without any previous user studies or clinical contacts this study was conducted with healthy older participants. It was expected that their computer experience and overall performance most closely resemble the characteristics of the target population with neurological disorders such as stroke.

All subsequent experiments were conducted at the Neurological Department of the Asklepios Rehabilitation Clinic in Schaufling, Germany. The pointing task and spatial memory task (Virtual Memory Task – VMT) were implemented in a virtual replica of the experimenter's office. Both tasks were tested with a wide range of neurological patients to determine whether task difficulty, user interface and psychometric properties were adequate for clinical use. In addition, the VMT was evaluated in a single case study in which a patient with severe traumatic brain injury used the task as part of her inpatient rehabilitation.

Lastly, the following aims and hypotheses were derived with the goal to demonstrate that the proposed system is appropriate for clinical use. Specifically, the aims and hypotheses of this dissertation are twofold, targeting clinical and technology aspects of system development.

1.8 Aims

I. It is the primary aim of this thesis to develop a set of cognitive tasks targeting navigation ability, orientation ability, and spatial memory.

II. It is an additional aim to assess each cognitive task's validity in an experimental trial.

III. It is aimed to develop an optimized workflow for creating individualized VEs.

IV. It is an aim of this thesis to integrate each cognitive task in a meaningful VE.

V. It is an aim to test the efficiency of the development process of the VEs:

   a. in a controlled setting.

   b. in a clinical, patient-centered setting.

VI. It is aimed to apply the embedded cognitive tasks throughout the neurological rehabilitation of a brain-injured patient.

VII. It is an aim to use the VMT to accommodate a patient's individual therapy goal.

VIII. It is an aim to integrate the proposed workflow into the rehabilitation routine of a brain-injured patient.

1.9 Hypotheses

I. Cognitive tasks integrated into the VEs are expected to target specific cognitive processes (process-specificity).
   a. The VNT is hypothesized to show equivalent outcomes of navigation measures as compared to a real-world navigation task.
   b. The VNT is predicted to significantly correlate with pencil and paper measures of spatial abilities.
   c. The VMT is hypothesized to significantly correlate with established neuropsychological tests that assess spatial memory.
   d. The VMT is hypothesized to significantly correlate with established neuropsychological tests that assess spatial abilities.
   e. It is predicted that the VMT does not show significant correlations with cognitive tests of domains unrelated to the VMT.
   f. The pointing task is hypothesized to show equivalent results in a real environment and its virtual counterpart.

II. The proposed applications are predicted to be flexible enough to meet the changing demands of a patient's neurological rehabilitation (context-sensitivity).
   a. The VMT is expected to be used throughout a patient's neurological rehabilitation without the occurrence of a floor or ceiling effect.

III. The workflow for creating the proposed individualized training is expected to be suitable to create realistic, high-fidelity environments with a high degree of ecological validity.
   a. It is hypothesized that developed VEs show high recognition rates by users.
   b. Cognitive tasks are expected to be transparent and easy to understand by users.

IV. The workflow for creating the proposed individualized training is expected to be effective enough for integration into the daily routine of a rehabilitation clinic.
   a. Each functional training environment should be created in less than one working day (i.e. eight hours of development).

2 SYSTEM AND PROCESS DESIGN

Development of clinical virtual reality (VR) systems requires the consideration of technological and clinical system requirements. Both aspects have to be taken into account when software and hardware choices for the application development are made. Many software packages are available to create interactive virtual content. Often, several programs have to be used in a more complex workflow to achieve best results. The requirements and available options which have been considered for the proposed VR application are presented in the following chapter.

The goal of the development effort is the creation of realistic virtual environments (VEs) that closely resemble real-world places. For this purpose high-quality, low-polygon 3D models are needed. Further, a set of cognitive tasks needs to be developed that can easily be integrated into the created VEs. Game or simulation engines can be used to create interactive applications which combine the VEs, cognitive tasks, a user interface as well as data recording. Realistic environments in the larger context of a simulation were chosen over the development of a game, even though identical software and hardware is required for either development effort. It was expected that the realistic nature of the simulation enhances the transfer of cognitive abilities to the real-world, even though this aspect was not evaluated during the course of this thesis.

2.1 3D modeling

Four applications have been evaluated for their use in rapid modeling of VEs. Applications were chosen for the amount of information available online as well as prior experiences of the developer. Table 1 presents the requirements for 3D modeling and how they are met for each of these applications. Requirements were selected based on estimated use for the development of VEs and were subjectively chosen and evaluated by the developer. Information was gathered through completing online tutorials, reading online forums and website descriptions of the respective applications as well as using trial versions of each of the tools.

Table 1. Requirements for 3D modeling software

Software / Requirements	Autodesk 3DS Max/Maya	Blender	Google SketchUp
Ease of use	+ many tutorials available - complex interface	+ improved interface since version 2.5 + many tutorials available - complex interface	+ very easy interface + less tutorials needed - less tutorials available
File Import/Export	+ all popular formats supported	+ all popular formats supported	+ all popular formats supported (Pro-Version)
Free resources	+ websites with free models	+ websites with free models	+ websites with free models + Google Warehouse
Texturing and UV-mapping	+ fully supported	+ fully supported	- limited support
Low-polygon modeling	+ full control over polygon count	+ full control over polygon count	- limited control over polygon count

2.1.1 Ease of use

There is a wide variety of 3D modeling software available, each with its own set of advantages and limitations. For the purpose of this dissertation, several software packages have been evaluated. Most important criterion is the rapid development of non-organic objects (i.e. architecture) with an easy entry to proficiently creating the required objects. User discussion forums as well as written and video-based tutorials (e.g. YouTube, www.lynda.com) are offered to learn the respective programs. Autodesk's Maya and 3DS Max[10] appear to fulfill both criteria while at the same time having many online tutorials and learning material available. Alternative programs of about similar quality and complexity (e.g. Modo, Lightwave 3D, Cinema 4D) do not offer a comparable user community and as much training material. Blender[11] as an open-source project is very attractive due to its free

---

[10] Autodesk Maya / 3DS Max – http://www.autodesk.com

[11] Blender – http://www.blender.org

availability and the large amount of training content. However, before Blender version 2.5 was published in its various development stages between 2009 and 2011, the program's interface was complex and unintuitive (Reynish, 2008). Lastly, Google SketchUp[12] appears to provide a unique approach to 3D modeling with a minimalistic interface. Architectural and interior modeling are considered to be SketchUp's main areas of application.

In summary, all of the tested programs provide the desired functionality so that even an inexperienced user can learn each of the applications with available tutorials and user forums.

2.1.2 Export and Import of popular file formats

File format compatibility between development applications is a basic requirement to form an effective workflow. 3D models need to be exported in formats that the subsequent applications can process. Further, high popularity of the native file formats of the 3D modeler are helpful for importing freely available models. The COLLADA file format .dae and Autodesk's .fbx appear to be the most common file formats supported by almost all 3D modeling programs. With the exception of Google SketchUp, all other modelers support both formats. Google SketchUp only provides extensive support for importing and exporting file formats, including the export of .dae and .fbx files, when the Pro version of the software is purchased.

2.1.3 Availability of free 3D models

Using freely available 3D models such as furniture and plants can reduce the development time and cost of environments. Many of the no-cost models on websites like www.turbosquid.com are of the .obj or .3ds format and all 3d modeling programs are able to import these free resources. In addition, Google SketchUp provides a large repository of 3D models via the Google 3D warehouse which allows cost-free objects to be directly imported into a 3D scene.

2.1.4 Support for texturing and UV-mapping

Realistic textures are a main contributor for the realism of a VE. All 3D modeling applications provide support for importing textures and applying them to 3D models via UV-mapping. Though, it is the complexity of this feature that differentiates the available 3D modeling applications. Without any experience in manually unwrapping meshes effectively and applying textures to them, SketchUp provides an easy entry to texturing 3D models in a short amount of time. Though control over the applied textures is fairly limited when compared to the complex texturing tools that the other 3D modeling applications provide. The choice of modeling application is a tradeoff between feature-sets and simplicity and should be made based on the developer’s preferences and skillset.

---

[12] Google SketchUp - http://sketchup.google.com/

2.1.5 Low-polygon modeling for real-time applications

When creating models for real-time 3D applications, optimization is important to maintain high frame rates while the application is executed. Computer games and simulations often apply a technique called low-polygon modeling to keep the system requirements for the end user's computer as low as possible. In contrast, animated movies, scientific and architectural simulations commonly use highly detailed geometry or complex models exported from CAD-applications that are not suitable for usage on average consumer hardware. Low-polygon models can achieve almost the same visual quality as their high-detail counterparts when techniques like bump-, normal- or parallax-mapping are applied. Consequently, a workflow for cost-effective applications includes precise control about the 3D model's number of polygons to avoid the need for expensive hardware to run the finished application. All 3D modeling programs with the exception of Google SketchUp allow the user to add, remove and edit polygons individually. Because of SketchUp's minimalistic interface, no exact control over the number of used polygons is possible. However, for the purpose of this dissertation and later use in hospitals all prototypical applications were run on high-end computers where a higher polygon count is of little consequence. During the course of 2010/2011 the target systems consisted of hexa-core CPUs and PCI-E 2.0 graphic cards with GDDR5 memory and DirectX11 support within a price range of USD 2000 to 2500. Given the rapid advances in computing hardware the relevance of polygon counts and optimization can change in the future.

Taking the listed requirements and the developer's preferences into account, Google SketchUp Pro has been chosen as the 3D modeling application. The program's minimalistic interface and the availability of free models via the Google 3D warehouse are the main aspects leading to this decision. However, for continued development and distribution of VEs to end users like hospitals or patients, the modeling workflow might require adjustment in the future to reduce the polygon count of the created models. This can either be achieved by using a third-party SketchUp plugin[13] that adds the desired functionality or by switching to one of the other available 3D modeling applications. All of the listed applications and others such as Luxology Modo[14] and Maxon Cinema4D[15] are viable solutions and each developer has a wide range of choices available depending on individual preferences.

---

[13] Artisan4SketchUp - http://artisan4sketchup.com/category/learn-more/

[14] Luxology Modo - http://www.luxology.com/

[15] Maxon Cinema4D - http://www.maxon.net/home.html

2.2 Task development and integration of VEs

Several game and simulation engines have undergone thorough testing regarding the suitability for the proposed software development. Vizard Lite Edition[16], Quest 3D VR Edition[17] and Unity Pro[18] have been used for projects related to the assessment of cognitive functions. Small prototypes were developed which included importing 3D models, implementing user input and data recording (Figure 1). These prototypes were mostly unrelated to this thesis and will not be discussed in detail. Based on subjective judgment, expected project tasks, and experiences gained from the development of the initial prototypes, a list of requirements was derived in order to select a game engine for this project. Information about each application was collected through user forums, documentation, tutorials and conversations with existing users. Additional programs such as UDK[19] and Torque3D[20] have also been evaluated, but have not been used more extensively. Their feature sets, target audience and business model did not seem appropriate for the purpose of this thesis. An overview of each application and how each addresses this thesis' requirements can be found in Table 2.

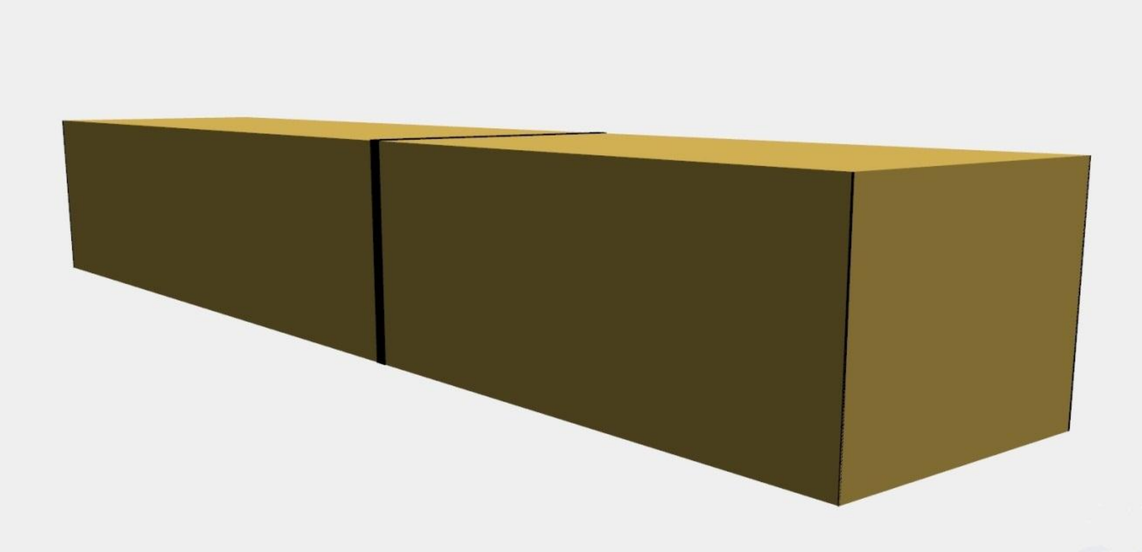

Figure 1. Prototype for Object Bisection developed in Quest3D

---

[16] Vizard VR Toolkit - http://www.worldviz.com

[17] Act3D - Quest3D – http://www.quest3d.com

[18] Unity Game Engine – http://www.unity3d.com

[19] Epic Games - Unreal Development Kit – http://udk.com

[20] Torque3D – http://www.garagegames.com

Table 2. Requirements for 3D game and simulation engines

Software / Requirements	Vizard VR Toolkit	Quest3D VR Edition	Unity Pro
Ease of use	+ uses open-source Python + many Python tutorials available - small user community	+ visual scripting + LUA support + helpful user community - lack of documentation	+ easy interface + many tutorials available + large user community
File Import/Export	+ most popular formats supported	- only Collada and .x format supported	+ all popular formats supported
Physics engine	+ is supported	+ is supported	+ is supported
Realistic rendering and lighting	- only basic support for lights and shaders	+ support of shaders, lightmaps, HDR lighting, shadow/normal maps - implementation is not user-friendly	+ Beast lightmapping implemented + support of post-processing effects, shaders, deferred rendering - most advanced features require Pro-Version
Support for input devices and VR hardware	+ support for many devices	+ support for many devices	- limited support through third-party developers
Royalty-free publishing	+ several product tiers available	+ supported	+ supported

2.2.1 Ease of use

Main requirement for choosing a development platform is the easy entry to developing and publishing interactive VEs. An active user-community, thorough documentation and the availability of online tutorials are of particular value for this requirement. Vizard has the advantage of using Phyton, a well-documented programming and scripting language for which there are many books and tutorials available.

Quest3D features a visual scripting approach supports development without any programming knowledge via node-based systems of logic blocks/channels. Custom logic can be integrated via the LUA scripting language and custom channels can be created via C++. However, the documentation of the latest release (Quest3D 4.x) appears to be inferior to other development engines, especially regarding the use of object-oriented development. On the contrary, the user community and forum have always been helpful and quick to answer questions.

Unity's game engine has been through tremendous changes since early 2009. After releasing a Windows-based version of the previously Mac-only engine in March 2009, Unity has seen a steady rise in its userbase. In October 2009, the former Indie-version of Unity was made available for free which resulted in another large increase in users. The growing community led Unity Technologies to extend their documentation, provide example projects and create a Question&Answer

website[21]. A large number of user-created tutorials and specialized courses provide detailed instructions for most of Unity's features.

2.2.2 Import of 3D models of popular file formats

All game and simulation engines have been chosen to support the initially selected file formats of 3D models. Vizard imports .obj and .3ds models that each 3D modeler is able to export. Quest3D has added support for the Collada format (.dae) during their Version 4.0 release which can be imported and exported in all tested 3D modeling applications. Unity is capable of importing most file formats supported by current 3D modeling programs. Autodesk's .fbx, Collada's .dae, Blender's .blend, modo's .lxo and many other file formats are supported. Further, Unity features an automatic update of 3D models whenever they are saved within the 3D modeling application. This results in a faster asset workflow between 3D modeler and Unity than any of the other tested engines. Unfortunately, Google SketchUp does not support saving files in .fbx formats directly so that each model has to be exported first before any changes are automatically reflected in Unity.

2.2.3 Physics-engine

Physics engines are integrated into the framework of game engines as so-called "middleware". While there are different physics engines with varying sets of features available, all of them are capable of basic collision detection to simulate realistic walkthroughs in VEs. Raycasts, basic simulation of gravity, forces, joints and rigidbodies are also commonly integrated in any physics engine. Advanced features such as soft body physics and destructible environments are not essential for the projects carried out for this thesis so that all game engines are generally capable of fulfilling the listed requirements.

2.2.4 Realistic rendering and lighting

A game engine's rendering and lighting systems are important for the realistic appearance of VEs. Without proper lighting and the effective use of a rendering engine, environments look flat and artificial. In addition, lighting and rendering environments are aspects of development that can have a large negative impact on the application's performance depending on the adequate use of lighting and rendering techniques. Vizard's rendering system is based on OpenGL including support for basic lights and shaders via the OpenGL Shader Language (GSLS). Quest3D provides a more extensive set of options including HSLS shaders, lightmaps, shadow maps, normal maps, HDR lighting and many more. Unfortunately, the documentation and implementation of these features are not suitable for new users. In order to achieve high-quality results, extensive experience with 3D graphics and shader programming are recommended. Lastly, Unity features forward and deferred rendering for better performance with a large

---

[21] UnityAnswers - http://answers.unity3d.com/

number of real-time lights. Shaders, Umbra's occlusion culling[22], Beast's lightmapping[23] and many post-processing effects are also included in Unity's functionality. Most importantly, many features are user-friendly for non-expert users as detailed documentation is provided and most effects simply need to be dragged and dropped into the virtual scene. However, several features (e.g. real-time shadows, render-to-texture, occlusion culling, etc.) are only available for Unity Pro users. Nonetheless, Unity's functionality and ease of use for producing realistically rendered environments are well-suited for the workflow of the proposed clinical application.

2.2.5 Integration of input devices and VR hardware

Communication with commercially available tracking systems, head-mounted displays, projection systems and input devices is either already integrated in game or simulation engines or needs to be added via the use of sockets or plugins. Vizard and Quest3D already offer built-in support for many of these systems so that no additional programming is required. Unity does not provide such support as it is mainly targeting game developers who often publish their games to mobile devices instead of visualization systems. However, Unity's large user community makes up for the lack of supported external hardware by developing plugins and third-party software which give access to such hardware and input devices.

2.2.6 Royalty-free publishing

While the publishing of finished applications is not the primary concern for experimental studies and prototypical development, long-term use of the individualized virtual rehabilitation concept has to be taken into account for choosing a development platform. Funding of future extensions of the proposed applications as well as commercialization through tech-transfer organizations have to be taken into account. Hence, commercial publishing of the finished applications as well as existing royalty schemes are relevant when choosing a development architecture. Worldviz provides three product tiers that are distinguished by different publishing options and feature sets. Vizard Lite Edition is neither viable for experiments nor end-user distribution because it only allows to publish branded non-fullscreen applications. Developer edition features royalty-free fullscreen publishing for non-commercial purposes. Enterprise edition adds royalty-free commercial distribution to the before-mentioned features. Quest3D features a similar tier structure for its products. However, all three editions (Creative, Power, VR) allow for royalty-free commercial publication of applications and are only differentiated by inclusion of features. Lastly, Unity allows for royalty-free commercial publishing with all of its products. Program versions are distinguished by platform (e.g. PC/Mac, iOS, Android) and feature set (Normal versus Pro). Given these publishing options for each development tool Unity and Quest3D provide better options by allowing for unrestricted publishing of all applications.

---

[22] Umbra Occlusion Culling - http://www.umbrasoftware.com/

[23] Autodesk Beast - http://gameware.autodesk.com/beast

2.3 Workflow

Based on the previously discussed software choices a workflow for the creation of VEs has been established. The workflow was optimized based on available resources and preferences of the developer. After testing different options of capturing the real environment (e.g. hand-drawn sketches, photographs and annotations on a smartphone and tablet) and replicating it as a 3D model the following workflow was chosen for the creation of individualized VEs. All choices were based on subjective judgment by the developer.

First prototypes of interactive 3D environments have been created with Quest3D Academic VR Edition (Versions 4.0 – 4.3.2). The visual programming concept and the licensing model were the main reasons for choosing Quest3D over Vizard. By 2008 no Windows-based version of Unity3D was available yet. Before any experiments were conducted in early 2009, the development platform was changed to Unity Pro for Windows, Version 2.5 (up to Version 3.3) to take advantage of Unity's user community and documentation. All finished environments are now enhanced by Unity-driven modular cognitive tasks. The applied workflow is depicted in Figure 2.

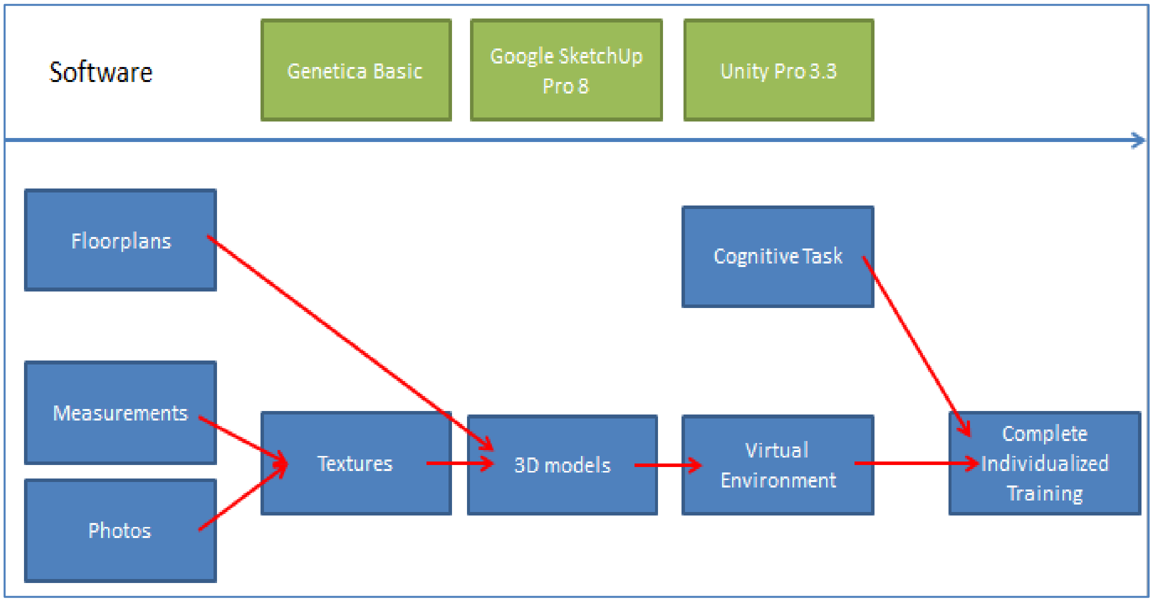

Figure 2. Workflow for development of individualized environments

2.3.1 3D models

Prior to modeling any 3D scenario, the real environment has to be measured and recorded. A 12-megapixel digital camera, a 6-meter tape-measure, a netbook, an HD USB-webcam, a Wacom digital tablet, the software Microsoft Paint and a trolley table to transport the equipment are used to gather all measurements. Photos of all localities are taken with the webcam. Measurements are added to each photo using the digital tablet and Microsoft Paint. This procedure appears to be the most effective, even when floorplans were available. Floorplans are often outdated and lack much information (e.g. heights, furniture, materials) so that photos have to be taken regardless of the availability of any other information. Finally, the digital camera is used to take high-resolution photos of materials and surfaces that are later turned into textures. The complete setup can be seen in Figure 3.

Figure 3. Recording station for photos and measurements

In order to create evenly focused, non-distorted textures these photos have to be taken at a 90-degree-angle towards the surface. For best results, flashlights and light reflections on the surfaces are avoided[24,25]. All photographs are imported into Genetica[26] and turned into non-distorted, seamless tileable textures. Texture resolutions are chosen depending on the texture's usage in the 3D environment. High resolutions (1024x1024 pixels or higher) are chosen for highly visible objects in the camera's foreground. Lower resolutions (e.g. 256x256 pixels) are more suitable for less exposed objects in the background. When no high-quality photographs are available, Genetica's texture generator or royalty-free textures from www.cgtextures.com are used. All 3D geometry is created and textured with SketchUp. Geometry is modeled for one-sided rendering so that unnecessary surfaces (e.g. inside of an object) are not rendered. Therefore, the direction of surface normals has to be taken into account when modeling. For easier handling of larger models, objects are separated in different layers depending on their category (e.g. windows, doors, furniture). Combinations of layers are then used to select, manipulate, export or hide certain geometry. Whenever possible, freely available models at Google's 3D warehouse[27] are used for furniture items. However, the

---

[24] 10 Texture Photography Tips - http://designm.ag/tutorials/photographing-textures/

[25] Ultimate Guide for Creating High Quality Textures - http://designm.ag/tutorials/photographing-textures/

[26] Genetica - http://www.spiralgraphics.biz

[27] Google Warehouse - http://sketchup.google.com/3dwarehouse/

theoretical advantage of having free models turned out to be of little use. Most available models are of too much detail (i.e. too many polygons) and mostly useless for real-time 3D applications. Reducing the free model's polygon count, cleaning up layers and aligning all surface normals often takes longer than modeling a complete object from start to finish. Completed 3D environments are exported as one-sided geometry from SketchUp via the .fbx file format. Larger models have to be split up into smaller parts to adhere to Unity's size limitation of 60.000 vertices per object. Reassembling those parts in Unity is not necessary as long as they were all exported within the same coordinate system, having the same pivot point. This is achieved by choosing "export selected geometry only" in SketchUp and selecting the appropriate parts of the model, preferably via the use of layers.

2.3.2 Scene setup in Unity

Once a model is imported into Unity, a second UV for lightmaps is created and mesh colliders are added to the geometry. The scene lighting consists of a combination of lightmaps and real-time lighting. If very few light sources are present and the scene is fairly small, only real-time lighting is used. In order to increase performance for large scenes with many lights, Unity's implementation of Beast lightmapping is applied to static lights and geometry. Unity’s in-built shaders are adjusted as needed, for example to display transparent materials like glass (diffuse/transparent shader). The use of collision geometry largely depends on the nature of the task. For tasks that allow unrestricted movement of the user it is necessary to create a simple version of the environment to simulate collisions. Detailed geometry like doorknobs and plants are replaced by simple cubes to avoid low frame rates whenever the user collides with these objects. The collision environments are imported separately into Unity and added as mesh colliders to the original environment. Tasks with minimal or no user movement simply use the original geometry as mesh collider. Lastly, asset packs that contain the cognitive tasks are added to the finished environments. The cognitive tasks are described in more detail in the following paragraphs.

2.4 Cognitive tasks

Three different cognitive tasks have been developed to be integrated into individualized VEs. Each task is designed to be simply added to an existing VE in the Unity editor. In order to be used in clinical sessions, the tasks have to be configured for the individual patient. For example, relevant targets or locations in the environment are selected and added to the task. The tasks' aim is to assess and train cognitive skills like spatial orientation and spatial memory. Despite the individualized task context, the process-specific nature of each task and the consistency of task-related features across each individual user are expected to provide the means for collecting normative data.

At the current prototypical stage of development, task setup can only be done by the developer as it requires in-depth knowledge of the virtual scene and the Unity editor. The goal for future development is to be able to set targets for each task by simply selecting them from a list and choosing from several task-specific options outside of the Unity editor (i.e. within the standalone application). Once the task has been configured, it is published as an executable file that can be run on Windows PCs.

2.4.1 Virtual Navigation Task (VNT)

The VNT requires the user to traverse from a specified starting point to a target position. Each target position can serve as the starting point for the next route, thus combining several stages into a longer route. As such the task can be used to train or assess the user's knowledge of familiar environments. Also, the user can be familiarized with important locations in unfamiliar environments. Both cases are important for patients in cognitive rehabilitation who are unable to leave the hospital. The procedure has been used in similar or slightly modified form in several wayfinding experiments in the past ((Waller, 2000; Witmer, et al., 1996). The task can be set up in the Unity editor by declaring any game object as a target. The object needs to be tagged as 'target' and several target-relevant scripts need to be added so that a task manager recognizes the item as a target. Once all targets have been set up, the training or assessment can commence. Navigation through the VE is implemented through any standard input device such as mouse, keyboard or joystick. Input for movement and viewing direction can be mapped to any button, mouse or joystick movement depending on the patient's needs. Movement and viewing direction are unrestricted so that the user can freely walk through the VE. However, several mechanisms are implemented to restrict user interaction with the environment when simulator sickness becomes an issue. Walking speed and view rotation speed can be set to match the user's comfort level. Further, extreme camera rotations (i.e. looking straight up or down) can cause distortions to the rendered environment. Thus, camera rotation can separately be limited or disabled on each axis in the 3D coordinate system to avoid users looking up or down. Currently, the application's capability to provide instructions is limited to displaying the navigation target (see Figure 4). The experimenter has to give any additional instructions and provide feedback about the user's performance.

Find the Lecture Theatre

Figure 4. Navigation task with onscreen instructions

During traversal the user's position and rotation within the VE are recorded at specified intervals (e.g. once every 100ms). All of these saved locations are written to a text-file and saved to the hard drive of the local computer. Further, the user's walked distance and time to reach the target are written to the text file. Once the session is finished, the text-file can be loaded in the Unity editor to display the locations in the VE. This allows the experimenter to plot the course of the user and measure distances and deviations between different routes.

2.4.2 Pointing Task

Pointing towards unseen targets requires the user to possess a mental representation of the environment. This technique was first used by Curtis, Siegel and Furlong (1981) and thereafter in several instances (Waller, 2000) to assess configurational knowledge about environments. The task places the user at a stationary location in the VE. From this position the user has to point towards predefined target locations. As the task is aiming to test mental representations of environments, the targets are occluded so that the user has no direct line of sight to each location. A red marker is displayed in the center of the screen. The user's task is to rotate the viewing direction in order to match the red marker with the target's exact location. Viewpoint direction can be manipulated by mouse, keyboard or joystick input. The application displays the name of the current target location onscreen. The red marker is explained by the experimenter to be similar to the user's index finger for pointing towards a target (see Figure 5).

Figure 5. Pointing task within a virtual model of a rehabilitation clinic

Users are told to point through walls, floors, ceilings or any obstacles in order to indicate the target's location. The users are given the chance to familiarize themselves with the environment before the task begins. Though, users are not allowed to walk around the environment at any stage before or during the task. Verbal feedback about user performance is given after all targets have been finished. After each target the experimenter has the option of testing the user's distance perception. Distance and height difference towards the target can be entered in text boxes. Height difference is defined as the vertical distance between the target's floor and the current position's floor in the VE. The user's answers (direction vector towards the target, distance to target, height difference) are written to a text file and saved to the hard drive of the local computer. The application automatically calculates the differences between the user's answers and the angular and distance measures towards the actual target locations. Setting up the task in the Unity editor is similar to the navigation task. Target objects are manually defined by the experimenter. Target-relevant scripts are added to the game objects. Once the task is set up, the application is published as an executable to be run on any Windows-based PC.

2.4.3 Virtual Memory Task (VMT)

The VMT combines elements of traditional working memory tasks (Kessels, Zandvoort, Postma, Kappelle, & Haan, 2000) and perspective taking tests (King, Burgess, Hartley, Vargha-Khadem, & O'Keefe, 2002). Its rationale has been inspired by experiments of King and colleagues (2002) and Shrager and colleagues (2006) in which VR tasks were used to assess spatial memory of brain-injured individuals. The task creates realistic scenarios for training and assessment of memory functions where information acquisition and retrieval are not necessarily identical (i.e. spatially rotated). The application can either be used to assess users in familiar and unfamiliar environments or to practice memory-related strategies by having the user find or place targets in specific locations in the environment. Task difficulty can be manipulated through the number of memorized target items and the angular displacement of the user's viewpoint. Both of the task's difficulty parameters, item number and angular displacement respectively, can be changed independently. The user is set at a stationary point in the VE and presented several target objects. The scene camera can be rotated freely but no movement is possible. The user has to memorize the exact locations of all the targets as accurately as possible. After a set amount of time the targets are moved to different locations in the environment. In addition, the scene camera can also be moved to a different position in order to initiate a viewpoint change. The locations of the target objects and scene camera can either be determined by the experimenter or randomly changed according to set parameters (e.g. randomly placed in a circular area in the environment). The randomized locations provide a variety of options for the application’s use as a long-term cognitive training. The predetermined locations are useful for assessments when several sessions have to be compared over time or between users. As soon as all targets have been moved the user has to drag the target objects back to their original positions. Items can be selected and moved through mouse or keyboard input. The order in which targets are moved is not relevant. When the user indicates that all objects have been moved to the correct positions, the experimenter has the option to give visual feedback by overlaying transparent markers of the original positions in the environment (see Figure 6).

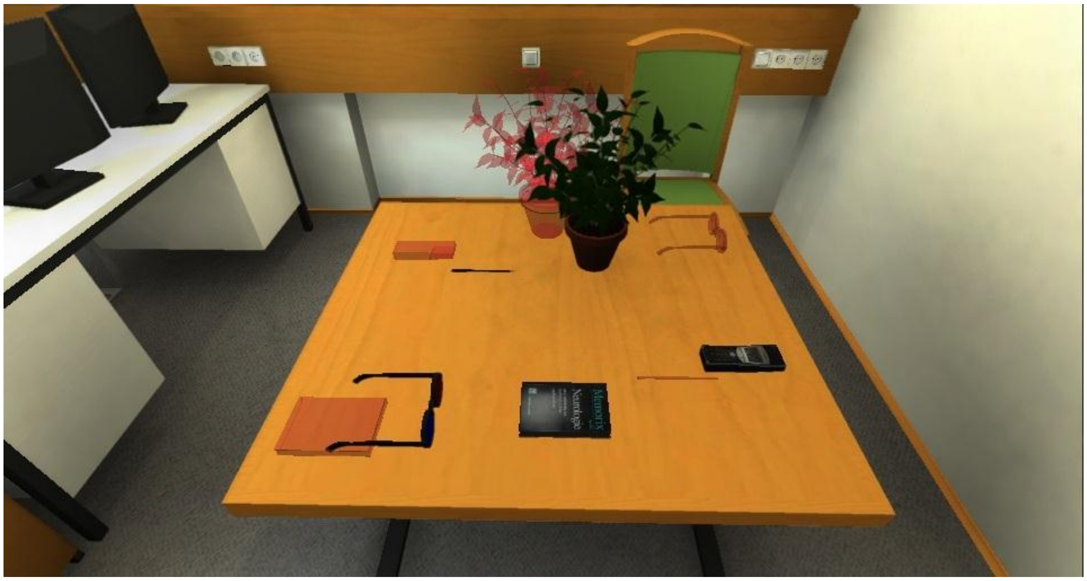

Figure 6. VMT with visual feedback of correct item locations

The application saves the positions of all objects at the different task stages to a text file. Distances are calculated between the original positions and the user's answers. All locations can be loaded back into the Unity editor to allow for post-session analysis. The Unity editor is also used to set up the task prior to a training session. Each game object in a VE can be used as a target by adding target-specific scripts to the object. Targets and their location before and after viewpoint changes are saved via a customized editor interface. A series of different tasks can be saved for each patient's training session and published as an executable file via the Unity game engine.

2.5 Costs

Based on the software choices and previously described workflow, the estimated costs for initial and ongoing development are listed in Table 3 below. It can be argued that Unity Free is a sufficient alternative if visual features such as real-time shadows and post-processing effects are of no importance. The actual costs for developing the proposed cognitive tasks are not included in the table. For the purpose of this cost evaluation the one-off cost of the task development is of no relevance. The proposed workflow and framework are intended for continuous use in a clinical context in which ongoing costs (i.e. development of individualized environments) are the main concern. Comparing the total hardware/software costs to sophisticated VR systems (e.g. Worldviz[28]), it becomes obvious that the proposed development workflow and delivery method (i.e. flatscreen desktop system) provide a cost-effective alternative to traditional VR setups. The workflow is based on the preferences of each developer so that prices for software and hardware can vary considerably. For example, Autodesk 3DS Max and Maya (instead of Google SketchUp) cost USD 3,999.00 and USD 4,090.00 respectively[29].

Table 3. Cost overview for hardware, software and development

Hardware / Software	Work Hours	Costs
Google SketchUp Pro		USD 495
Unity Pro/Free		USD 1500 / Free
Autodesk Design Review		Free
Genetica Basic		USD 149
Misc. equipment (digital camera, measuring tape)		~USD 150
Windows PC incl. screen, mouse, and keyboard		~ USD 1500
Total hardware/software costs		USD 2294 - 3794
Initial task development	1 developer x approx. 2000 hours	varies according to hourly rate
Ongoing development per patient	1 developer x approx. 8 hours per room	varies according to hourly rate

---

[28] Worldviz VR hardware/software: http://www.worldviz.com/purchase/pricelist.php

[29] Autodesk Online Store: http://store.autodesk.com/DRHM/store

2.6 Summary

Taken together, the above mentioned components of the proposed application form the basis of a clinical tool which combines context-sensitive and process-specific features. Firstly, the efficient workflow is expected to provide the means for an expert developer to create individualized VEs for each patient. It is hypothesized that the patient is able to recognize the VE and train in task context that is similar to relevant real-world scenarios. Despite the effort to build new 3D models for each VE, this context-sensitive setup has to be efficient enough for integration into a busy routine at a rehabilitation clinic. Secondly, the cognitive tasks need to provide content for neuropsychological training and assessment which is specific to distinct cognitive functions and validated through validity analyses. Through the modularization of the cognitive tasks, the proposed application forms a symbiosis of context-sensitive and process-specific aspects of cognitive rehabilitation. The tasks are expected to be re-used across patients while the context is rapidly created for each brain-injured individual. Detailed analyses of the application's validity and clinical usage are discussed in the next chapters.

3 SYSTEM EVALUATION

The following chapter describes the evaluation and validity analyses of the software that has been developed as part of this thesis. The chapters will be presented in the order that their respective trials have been conducted. Visual quality and functionality of the developed virtual environments (VE) gradually advanced over the course of the four experimental trials. The Virtual Memory Task integrated in a complex VE (Experiment 1) was the first functional prototype of this thesis to be systematically tested with participants. Given the prototypical nature of the assessment and the small number of available brain-injured patients in Christchurch, New Zealand, healthy adults were recruited for this study. All other experimental trials were conducted with brain-injured individuals at the Asklepios Rehabilitation Clinic in Schaufling, Germany. The clinical trials were conducted during a six month overseas visit in Germany between October 2010 and April 2011.

Usability and accessibility testing have been part of the development cycle of each of the study's prototypical applications. However, most user feedback has been collected verbally through unstructured interviews or open discussion with the participants. The feedback often was integrated into the applications straight away or before the next test session was conducted. Hence, the majority of feedback and changes has not been documented. Whatever documentation of user feedback exists will be integrated into the respective chapters.

Prior to the development of the Unity-based applications that have been used in this dissertation's experiments, several simple prototypes have been created using the Quest3D simulation engine. Even though these early programs were valuable for developing a workflow for VEs, none of them have actually been used in any data-collection and hence will not be mentioned in any detail here. However, these applications lay the foundation for the data recording which has been used throughout this dissertation's experimental trials. The procedure consists of repeatedly writing the user's position and rotation within the VE to a text-file on the local computer (e.g. once every 100ms). These position and rotation vectors can then be used to plot the user's movement in the actual environment by reading them from the text-file and feeding them back into the application. At each saved location a simple primitive is placed in order to visualize the user's path through the environment (see Figure 7 and Figure 8).

Figure 7. Early Quest3D-prototype of navigation analysis

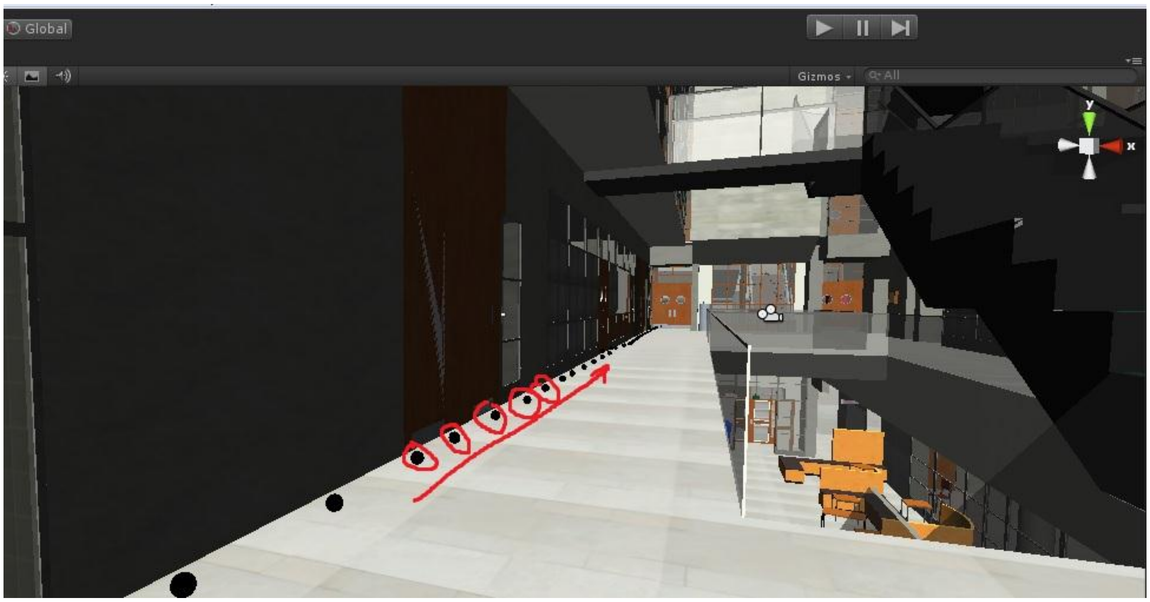

Figure 8. Analysis of traveled path in the Unity editor (path annotated in red)

A similar approach is used for saving other task-related parameters such as traveled distance or time-to-finish-task. The resulting text-files are arranged so that they can easily be imported into Microsoft Excel or SPSS for further analyses.

3.1 Experiment 1 - Evaluation of the Virtual Navigation Task

This experimental study has been carried out in early 2010 at the University of Canterbury, Christchurch, New Zealand. Results of this evaluation have been published in the proceedings of the International Conference on Disability, Virtual Reality and Associated Technologies 2010 in Vina del Mar, Chile (Koenig, Crucian, Dalrymple-Alford, & Dünser, 2010). Further, the results have been published in the International Journal on Disability and Human Development (Koenig, Crucian, Dalrymple-Alford, & Dünser, 2011). The aim of this study is to validate the previously described VNT. Target locations can be created within any VE and navigation performance between these locations is recorded for post-hoc analysis within the game engine Unity. This approach is the basis for assessing large-scale spatial abilities in patients with brain injuries as part of their context-sensitive rehabilitation. Specifically, this experiment aims to assess whether large-scale navigation in a real and identical VE are equivalent. This approach is different to previous experiments in which transfer between virtual and real environments were the primary focus (Waller, 2000; Witmer, et al., 1996). Transfer in such experiments refers to the improvement of performance in the real environment when the user was exposed to a virtual model of the environment. This improvement can simply mean that users find their way through the real environment regardless of how this outcome is achieved. However, no equivalence can be inferred as it was not analyzed how performance in both environments compares in regards to the participants' behavior. Hence, analyses of transfer of knowledge and comparison of equivalence of performance in two related environments (virtual and real) are conceptually different. Equivalence evaluation requires the analysis of how participants navigate in both environments and whether similar strategies or errors occur. This experiment attempts to explore such behavior to go beyond existing research findings of transfer of wayfinding knowledge (Standen, Brown, & Cromby, 2001; Witmer, et al., 1996).

3.1.1 Introduction

Navigation is a highly complex skill of moving oneself, a craft or vehicle through novel and familiar environments. For many brain-injured patients with cognitive deficits the ability to return back home or to work are primary goals. Therefore, real-world navigation through complex familiar and semi-familiar environments (e.g. neighborhood, route to work) is a key element of independent living. A variety of cognitive functions such as memory, visual and spatial perception and problem solving are involved in navigating through such environments. If any of the involved cognitive functions is affected by a brain injury, the amelioration of navigation deficits is an important part of cognitive rehabilitation. However, navigation training during rehabilitation is often restricted to very few locations like the hospital or the patient's home. When faced with such limitations it is desirable to use simulations to retrain patients' lost abilities in a wide range of environments. VEs with high ecological validity can provide the means for individualized, context-sensitive training of navigation abilities. However, before such technology can be applied, it is necessary to evaluate how real-world navigation and virtual-world navigation compare, especially when complex environments are of interest.

Until now, knowledge transfer between real and VEs and differences between several modes of knowledge acquisition have been studied. The results of these studies have been mixed. In a study by Richardson, Montello and Hegarty (1997) healthy participants learned the layout of a complex building either from a map, the real building or a VE similar to the real building. Test performance in the real building yielded significantly poorer performance by the VE group. After multiple training sessions within a large virtual building Ruddle, Payne and Jones (1997) were able to demonstrate near-perfect route finding abilities of their participants. Koh, von Wiegand, Garnett, Durlach and Shinn-Cunningham (1999) compared real-world training with the participants' exposure to immersive, non-immersive visualizations and also an architectural 3D model of the same environment. During training participants were free to explore the environment. While the authors concluded that training in virtual and real space are comparable, no actual navigation behavior was required during the testing phase and only estimations of bearings and distances were reported. Taken together, many navigation studies have been limited in several ways. They only involve learning a single predefined route or judgments of bearings and distances from stationary viewpoints. This type of learning is valuable when demonstrating training effects from simple VEs to the real world, but is not sufficient to specify how people navigate through their surroundings. It is also inappropriate for making predictions of real-world navigation behavior which is desirable in a clinical context. Looking at predefined routes or knowledge of landmarks poses obvious restrictions compared to finding your way through a complex environment. Moreover, when people navigate in their daily life, their goals and priorities change often and unforeseen circumstances and obstacles arise, so that a single predefined route is not always a viable solution. Routes cannot always be rehearsed in advance and the navigator has to make inferences about alternative routes and the overall spatial layout of the environment. Assessing such configurational knowledge about the environment in addition to route knowledge is a step in the right direction. Witmer, Bailey, Knerr and Parsons (1996) trained their participants in a complex office building and assessed route and configurational knowledge. However, their study is still limited to a predefined route and landmarks along that route. The authors' results suggest that using VEs for route learning is superior to maps, but inferior to real-world training. Examining navigation behavior in all its complexity, this present study explicitly compared human navigation in a large real-world building and its virtual counterpart. The measurement of navigation behavior is part of this thesis' proposed framework for assessment and training of cognitive skills in a clinical context, with focus on patients with brain injuries. With such focus it becomes important to assess how and why people are getting lost. Thus, an important aspect for this study's design is the high demand which is placed on the participants' navigation skills to provoke situations of temporary disorientation. The developed virtual reality (VR) simulation is intended for use in the day-to-day routine in rehabilitation settings. As such, usability, flexibility and compatibility with the needs of brain-injured individuals are of highest importance. It is the aim of this application to assess large-scale navigation ability and to make predictions about navigation performance in the equivalent real-world environment. Such predictions require that a VR-based simulation evokes similar behavior as compared to a real-world scenario.

This experiment is intended to test the hypothesis that critical navigation parameters are equivalent in a VE and its real-world counterpart. Specifically, walked distance, number of received cues, number of decision errors at intersections, distance estimations and number of pointing errors are expected to be equivalent in both environments. Navigation time, number of stops a participant makes and total time spent standing still during navigation are predicted to be higher in the VE, as these variables are expected to be influenced by the interface of the computer application.

3.1.2 Methods

3.1.2.1 Participants

36 healthy, right-handed participants from the Christchurch community aged 40 or older and unfamiliar with the tested building volunteered for this study. Only 29 participants are included in the analyses as three participants withdrew from the study due to symptoms of simulator sickness, two participants were familiar with the tested environment and two participants were excluded due to missing data after a technical failure of a recording device. The specific age group was chosen to include users with a wide range of computer experience and to assess the age bracket of patients with higher chances of stroke who are expected to be a primary target group in the future. Potential availability of patients with brain injuries was discussed with clinicians from collaborating hospitals and brain injury community groups in Christchurch. The issues of patient dropout, informed consent, comorbid conditions (e.g. physical disabilities, depression, aphasia) and safety risks (walking around a busy campus, simulator sickness) for patients was considered during these discussions. Also, feasibility of recruiting brain-injured patients and bringing them to the university campus for a three-hour test session was judged to be problematic. Consequently, healthy adults were chosen instead of an acute or chronic clinical sample.

Age of the participants ranged from 51 to 72 years in the real-building group while the age for the VE group ranged from 42 to 66 years. Male and female participants were equally assigned to both groups – six male and nine female participants in the real-building group and five male and nine female participants in the VE group.

3.1.2.2 Design

Participants were assigned to either a real-world or VE group in a randomized blocked design. Each participant was shown the same set of 12 target locations within the real version of a complex building on campus of the University of Canterbury, New Zealand. Following the initial learning phase, a series of pen and paper tasks for assessment of spatial abilities were completed. Finally, half of the participants returned to the real building (real-building group) to find the previously shown locations while the other half (VE group) was asked to complete the same tasks in the virtual version of the campus building. No follow-up assessment was included in the study protocol.

3.1.2.3 Materials

3.1.2.3.1 Real and Virtual Environment

The assessed environment was the seven-floor Erskine Building at the University of Canterbury, New Zealand (Figure 9).

Figure 9. Erskine building, University of Canterbury

The building's lower four floors were chosen for their complexity and unusual layout. Several staircases throughout the building allowed for a large amount of possibilities to traverse from one landmark to the next. An example of one floor plan can be seen in Figure 10. The virtual model of the building was created using Google SketchUp 7 Pro (see Figure 11). Textures were imported from photographs and floor plans were used to model the building to scale. Floor plans and measurements were displayed with Autodesk Design Review 2011. Interactions within the VE, data collection, interface and visual and navigation analysis tools were developed with the game engine Unity (version 2.6).

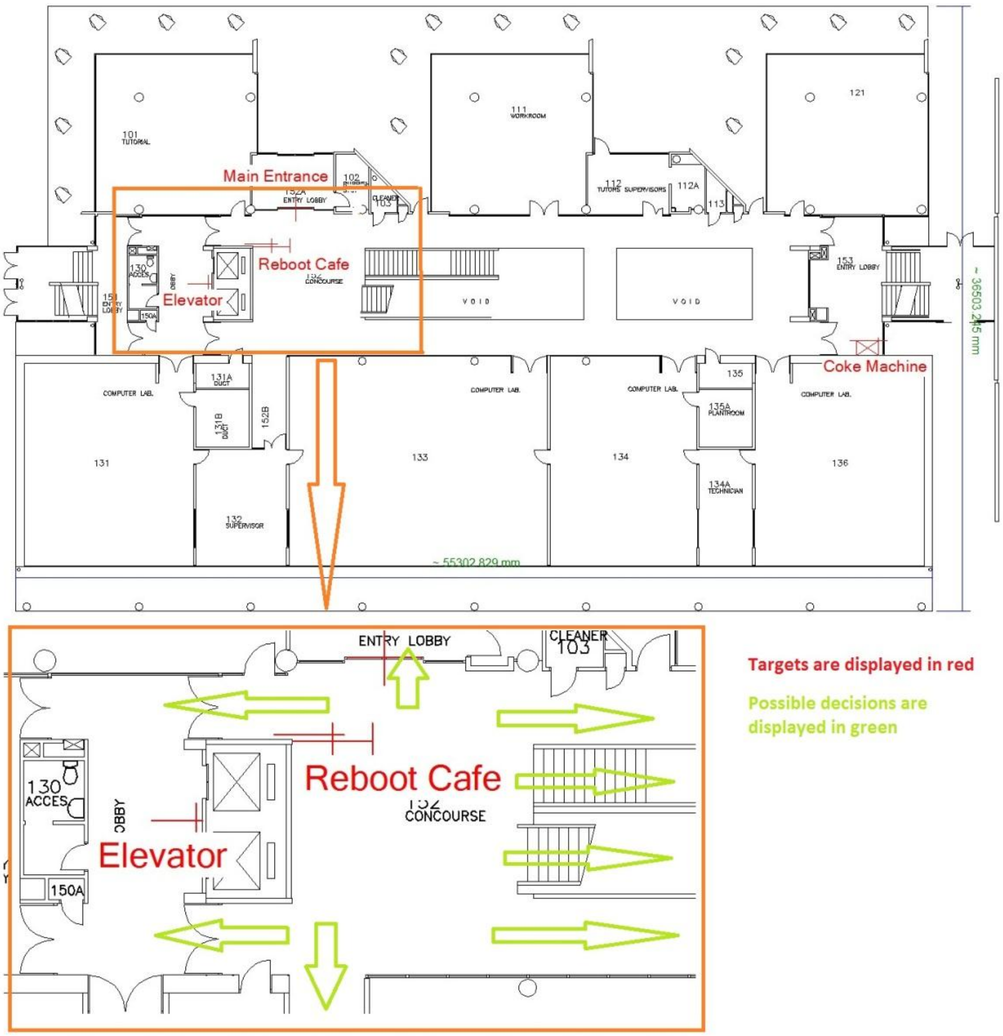

Figure 10. Floor plan of the Erskine building's ground floor

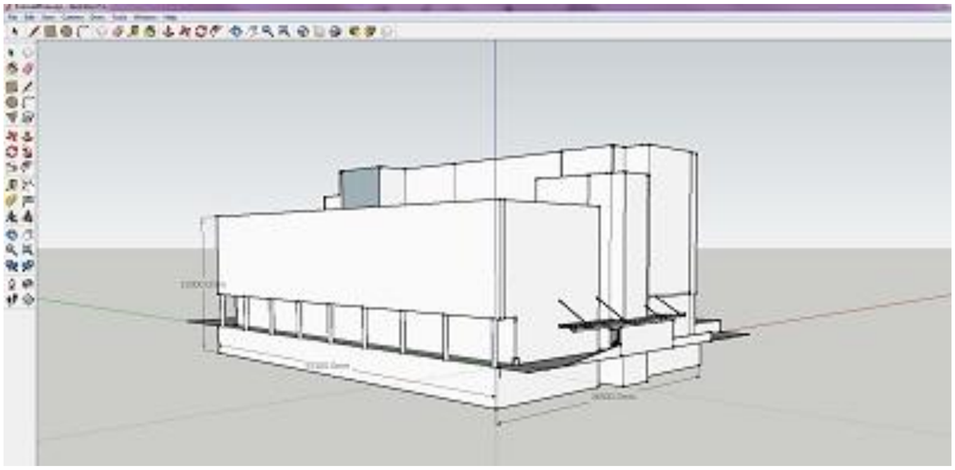

Figure 11. Model of the Erskine building in Google SketchUp

The VE was displayed using a three-screen back projection system with a field of view of 120° (see Figure 12). However, due to technical limitations the displayed VE appeared slightly stretched at the left and right edge of the screen. Each screen measured 2.44m x 1.83m. The participant was seated 2.2m in front of the center screen. This set up allowed the participants to show natural orientation behavior by turning to the side screens for searching the environment. The VE was rendered using a quad-core PC with three Nvidia GeForce GTX260 graphics cards running in SLI and a Matrox TripleHead2Go graphics expansion module. Participants were provided a standard three-button computer mouse to navigate through the environment. Cost-effectiveness of the used projection setup was only a minor concern for this initial prototype. The VisionSpace Theater was chosen over LCD screens and a head-mounted display for reasons of user comfort, field-of-view, Unity integration and availability. Results from Bowman and colleagues (2002) as well as Santos and colleagues (2009) suggest that the choice of display solution has an influence on user performance. Clear relationships with computer experience, input device and demographic characteristics of the studied population have not been established yet. In future studies head-mounted displays and desktop setups with LCD screens should also be taken into consideration.

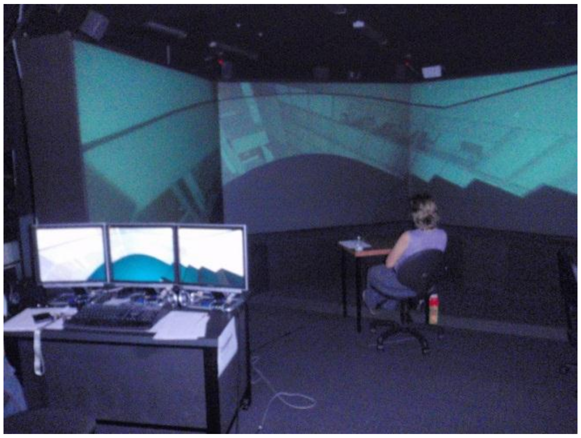

Figure 12. VisionSpace Theater, HIT Lab New Zealand

3.1.2.3.2 Pen and paper tests

Spatial abilities were measured with the Object Perspective Taking Test (OPTT) (Hegarty & Waller, 2004), Mental Rotations Test (MRT) (Vandenberg & Kuse, 1978) and the Card Rotations Test (CRT) (Ekstrom, French, Harman, & Dermen, 1976). In addition, orientation ability was assessed with the Santa Barbara Sense of Direction Scale (SBSODS) (Hegarty, Richardson, Montello, Lovelace, & Subbiah, 2002). Simulator sickness was assessed using the Simulator Sickness Questionnaire (Kennedy, Drexler, Berbaum, & Lilienthal, 1993). Computer experience was measured with an adapted version of the Computer/Internet Experience and Skills Questionnaire for: Internet Diabetes Trial at Harborview (Goldberg, 2006).

The OPTT requires the participant to judge bearings from imagined viewpoints which are not aligned with the participant's viewpoint. Each judgment is compared against an angle which is defined by a constellation of three objects out of an array of seven objects drawn on a sheet of paper. The average judgment error is calculated for the absolute angular deviations across the test's twelve items.

The MRT and CRT require the correct identification of test objects in comparison to target objects. Test and target objects are three-dimensional line drawings for the MRT and random two-dimensional polygons for the CRT. The number of attempted test objects divided by the number of correctly identified test objects is used as test score. The SSQ is a self-report measure for severity of simulator sickness symptoms.

3.1.2.3.3 Navigation Test

Navigation through the Erskine Building consisted of two phases. During an initial learning phase, all participants were guided through the building on a predefined path which passed 12 target locations on four different floors. The total length of the learning route was 498 meters. The lower four floors contained a total of 26 decision points where participants had to choose between alternate paths.

Alternate paths were classified as optimal, suboptimal or wrong. The optimal path was defined as the shortest single route which takes the participant from start to target. A wrong path is a decision which leads towards a wrong floor (i.e. target is up but participant goes down), along a route which does not lead to the target at all (i.e. dead end or wrong room) or any decision which is a direct turnaround on a path which leads optimally or sub-optimally to the target. All chaotic movement which cannot be classified as walking a defined path was considered a wrong decision. All remaining path choices were evaluated for the travel distance they require to reach the target, assuming that all following choices minimize travel distance. The shortest of these paths is suboptimal, all others are wrong. Optimal decisions were analyzed separately whereas suboptimal and wrong decisions were combined to an error score. Suboptimal decisions were scored as an error with a factor of one and wrong decisions were scored with a factor of two. The error score is the sum of all non-optimal decisions.

Half of the twelve target locations were secluded and allowed no direct line of sight to the other locations whereas the other half was in a more central location with higher visibility towards other locations and the layout of the building. However, in such a complex environment it was impossible to control the order and amount of exposure that each location received during the initial learning route. The hidden locations naturally received less exposure whereas the central locations were seen more often during route traversal. Before walking along the learning route the participant was instructed to pay attention to the target locations and more importantly, to get a good sense of the overall layout of the building. Instructions also included the fact that the traversed learning route and order of target exposure were irrelevant for the following navigation test. Further, it was mentioned that all target locations were again to be rehearsed before starting the navigation test. Participants had to stay within the line of sight of the experimenter at all times. Walking speed along the route was held constant. Orientation behavior was strictly encouraged and initial instructions emphasized that the participant was free to do what he/she normally does when being in a novel environment. For example, participants were allowed to stop, turn around and orient themselves as long as they kept close to the experimenter.

During the assessment phase of the experiment participants were expected to demonstrate configurational knowledge from the very beginning by starting from a different building entrance and finding new ways through the building. Half of the target locations were designated navigation targets while the other half was used for pointing tasks. Instructions, order and nature of tasks were the same for all participants and both groups. Navigation and pointing tasks were always alternated in a sequence which did not match the learning route.

For navigation tasks participants were instructed to find the shortest way to the given target without using elevators or asking people for help. There were no route restrictions and all of the first four floors were available for use. Cues were given systematically whenever participants asked for help or indicated that they were lost. Further, whenever a participant took more than two consequent wrong turns at a decision point or when no progress was made on a wrong floor (>4 decision points without leaving the floor towards the correct floor), a cue was given. Cues were categorized to either state that the participant is on the wrong floor, to verbally identify the correct floor, to give a semantic cue about the target, to guide towards the correct side of the building, and to explain in detail how to get to the target. Cues were given gradually in the listed order except when a participant asked specifically for a cue. Participants' navigation performance in the real building was recorded on video. All videos were later analyzed using VirtualDub[30] to extract the timing of all tasks, cues, stops and to plot the exact route on a floor plan using Autodesk Design Review 2011 (Autodesk Inc.). In addition, the plotted paths on the floor plans were rewalked in the VE with accurate timing to visualize and analyze the data. That is, by transferring all data into the VE, distances, number of stops, angles and viewpoints were easily computed and displayed in 3D space with pinpoint accuracy. For visual and computational analysis of the participants' routes, the original VE was modified using the Unity game engine to allow the experimenter to visualize all data, rewalk routes and carry out distance and angular calculations after the experiment was finished.

As soon as a navigation target was reached a tripod with an attached protractor was set up at a predefined pointing location. The tripod had a wooden plate mounted on top with an attached clock-hand. The protractor was hidden underneath the wooden plate to prevent giving any cues to the participant. The clock-hand was used by the participant to indicate the direction in which the pointing target was expected to be (see Figure 13).

---

[30] VirtualDub – www.virtualdub.org

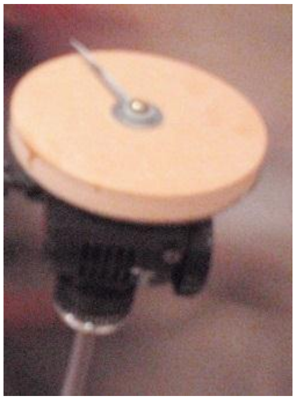

Figure 13. Pointing device

The absolute deviation from the correct angle was recorded. Pointing targets were always on the same floor and not visible from the participant's position. Participants were not allowed to leave the location where the tripod was set up. After pointing towards the pointing target, the participant was asked to estimate the egocentric euclidian distance towards the same target. The participant's answer was scored as a percentage of the actual target distance. Lastly, an empty floor plan (only the outer walls of the current floor of the building) was provided in which the participant had to draw his current and also the position of the pointing target. The location of two building entrances was shown on the floor plan to give the participant a better sense of distance and location. To analyze the participant's answers the floor plan was divided into a three by three array of sections. The deviation from the correct section was counted so that diagonal movement was not allowed. The highest possible error score for any response is therefore four, given that the participant's mark is in the opposite corner of the building from where the correct target location is.

3.1.2.4 Procedure

Participants were tested individually with each session lasting two to three hours. The experiment started at the Psychology Building at the University of Canterbury, New Zealand. After an initial briefing participants gave written informed consent. During the first 15 minutes questionnaires for demographic background, computer experience and the SBSODS were completed. Next, the participant was taken to the Erskine Building. The experimenter led the participant along the predefined learning route through the building and explained all twelve targets. After leaving the training environment, the participants returned to the Psychology Building where they completed the MRT, OPTT and CRT. Before navigation performance was tested, a list of the twelve targets was presented. Feedback and further explanations were given about all targets until the participant felt confident and no questions remained.

For the navigation assessment half of the participants were guided towards a different side entrance of the Erskine building. The other half of the participants was tested at the VisionSpace Theater of the HIT Lab New Zealand. A simple environment with two visible targets was used as a practice scenario for navigation and pointing tasks. After each participant was comfortable with using the mouse, the VNT started at the exact same virtual side entrance which was also used for the real-world assessment. The remainder of the testing session was identical for both groups as described previously.

3.1.2.5 Statistical Analysis

The basis of the present study is a comparison of real and virtual navigation with a prediction of equivalence being made between both groups. This by itself poses a problem, because traditional null hypothesis significance testing intends to test for differences between experimental groups/conditions. Further, the absence of a significant difference is not to be interpreted as both groups being equivalent (Nickerson, 2000). The research field of pharmacology, specifically pharmacokinetics, has made use of the two one-sided test as a method for equivalence testing (Schuirmann, 1987). Tryon (2001) and Tryon and Lewis (2008) suggest and alternative method through the use of a range (delta) which is defined by the extreme points of adjusted inferential confidence intervals (ICIs) of both groups. If delta is smaller than a predefined range of indifference, statistical equivalence is established. However, the range of indifference needs to be determined on substantive grounds which is not a trivial task for any research question. In conclusion, a hybrid approach to statistical analysis has been chosen. Firstly, both groups were compared with t-tests for independent means seeking to find a significant difference. Effect sizes for all comparisons were also calculated following the procedures of Cohen (1987). Lastly, the ICIs for both groups were calculated and their overlap was determined. The overlap measure indicates the percentage of participants which are included in the overlapping area of both groups' ICIs. A high overlap of the ICIs indicates that the results of both groups are similar, providing further evidence towards the equivalence of the experimental conditions.

3.1.3 Results

Levene's tests for homogeneity of variances were conducted for all comparisons of navigation performance and pen and paper tests. The variances for Total Time of Stops differed significantly between the VE and real building group ($F(1,27)=14.89$, $p=0.0006$). No other Levene's tests showed a significant difference. Lilliefors' tests for normality indicate that none of the distributions reported in this study differed significantly from a normal distribution.

As expected, most critical navigation parameters did not show a significant difference at $p = 0.05$ (see Table 4). Performance in the pen and paper tests was not significantly different between the real-world and VE group. There was no effect evident for the CRT ($d=0.19$). MRT, and OPTT showed small-medium effect sizes of $d=0.48$ and $d=0.41$ respectively.

For navigation distance no significant difference was observed. Participants in the VE group on average travelled 46 meters further than participants in the real environment. Cohen's d was found to be small for this comparison ($d=0.37$). Confirming our expectations, there was a significant difference between navigation time of both groups. Participants spent significantly more time navigating through the virtual Erskine Building which resulted in a large effect size (d=1.76). Variance for the VE group was large and individual performance ranged from 501s to 2111s. Navigation time was based on the total time for all six navigation tasks. Pointing tasks were not included in this measure.

The number of cues that were given to the participants did not differ significantly between both groups. However, a medium effect size (d=0.6) indicates that there was a difference in the number of help cues the VE group and the real-building group received even though this difference did not reach significance. This might also be due to the large variance in the VE group where two participants received 20 and 21 cues respectively. When both participants were excluded from the analysis, the effect size was reduced to d=0.24. A t-test for independent means showed no significant differences between both groups when the outliers were removed (t(25)=-0.622, p=0.539).

Table 4. Comparative measures, real and VE group

Measure	Mean+S.D. Real Group	Mean+S.D. VE Group	Df	p-value	Effect size d	CI overlap %
Decision Error Score	23.14+12.28	26.64+12.88	27	0.4684	0.28	21
Optimal Decision	12.86+2.96	13.36+2.76	27	0.6477	0.17	29
Number of Cues	4.87+3.68	7.86+6.32	27	0.1281	0.6	3
Floor Plan Errors	6.73+3.37	7.57+4.29	27	0.5620	0.22	38
Distance Estimation	120.01+50.91	90.42+38.4	26	0.0944	0.66	3
Angular Pointing Errors	30.98+19.35	45.23+19.85	27	0.0608	0.73	0
Navigation Distance	443.61+132.92	490.1+118.09	27	0.3296	0.37	24
Navigation Time	674.73+248.78	1380.5+520.63	27	0.0001*	1.76	0
Number of Stops	17.73+10.43	40.36+17.63	27	0.0002*	1.61	0
Total Time of Stops	171.4+111.62	589.75+388.33	27	0.0004*	1.67	0
CRT	0.89+0.087	0.91+0.057	27	0.6037	0.19	23
MRT	0.82+0.07	0.85+0.06	26	0.2231	0.48	13
OPTT	38.16+23.83	30.22+14.8	27	0.2953	0.41	6

Note: * indicates a significant difference at $p < 0.05$; CRT – Card Rotations Test; MRT – Mental Rotations Test; OPTT – Object Perspective Taking Test

The comparison of navigation decisions is of central importance, as this measure directly quantifies how participants navigated through the real and VE. As predicted, no significant differences were found for the number of optimal decisions and the decision error score. A small effect was found for the number of optimal decisions (d=0.17) and a small effect was evident for the decision error score (d=0.28). Systematic errors when drawing navigation and pointing targets onto empty floor plans showed no significant difference. The effect size for this comparison was found to be small (d=0.22). No significant difference was found for distance estimations in both groups. Nonetheless, participants who were assessed in the VE appeared to consistently underestimate the true distances towards pointing targets (d=0.66). A large effect size for angular pointing errors (d=0.73) indicates that pointing errors were larger in the VE than in the real building. The difference between both groups was non-significant. The remaining comparisons for number of stops and total time spent standing still both showed large effect sizes (d=1.61, d=1.67) and significant differences between both groups. Participants navigating through the VE stopped significantly more often and spent more time without virtual movement.

In addition to the aforementioned analyses, inferential confidence intervals (ICIs) were calculated for both groups of all measures (Tryon, 2001; Tryon & Lewis, 2008). A large overlap of ICIs, that is a high number of data points in the overlapping range of both ICIs, is an indication for the equivalence of both groups. However, as a result of this analysis almost no overlap was evident (see Table 4). Floor plan errors showed the highest overlap with eleven of the 29 participants in the overlapping range of the two groups' ICIs (38%). All remaining participants were located to the extreme left and right of the distribution of error scores.

Correlations of pen and paper tests (MRT, CRT, OPTT) with our navigation parameters were non-significant throughout. Only CRT score and errors in the floor plan task correlated significantly (r=-0.516, p=0.007) such that higher CRT scores were associated with less errors in this task. Also, age and gender showed no significant relationship towards any of the navigation measures.

Computer experience of the participants in the VE group was correlated with all navigation outcome measures. Correlations were generally negative and non-significant. Computer experience and the number of optimal decisions along the traversed route were positively correlated (r=0.564, p<0.05) which suggests that participants with higher computer experience performed better in the VE.

The participants' experience with the VE was almost entirely positive. Participants were asked to report symptoms of simulator sickness at the start and end of the test session as well as during a debriefing period after the session. Few participants reported mild symptoms of simulator sickness and three participants had to withdraw from the study due to more severe symptoms. The average increase of the total score from pre-assessment to post-assessment was 32.21 (SD=40.37) over 18 participants.

3.1.4 Discussion

In this experimental trial, validity of behavioral measures in a complex building was assessed and navigation performance in the VE and its real-world counterpart were directly compared. The VNT focused on configurational knowledge of the building and the 29 participants were required to make inferences about the shortest routes which had not been part of their previously shown learning route.

Most navigation parameters did not show a significant difference between the real-building and VE group. When participants were required to make decisions along their travelled routes, their decision errors and choices for the optimal, shortest route did not differ significantly between groups. In addition to the standard statistical analyses, effect sizes were calculated in order to further support the hypotheses of equivalence of both groups. Small effect sizes for both variables supported the initial null hypothesis tests. Another variable of importance is the number of cues which the participants received to find the navigation targets. Again, both groups did not differ significantly and a medium effect size was observed. Only after removing data of two participants who received most of the cues in the VE group, the effect was reduced to small size. Both participants had difficulties adjusting to the VE and using the navigational interface. Due to their difficulties to navigate adequately, abrupt viewpoint changes resulted in symptoms of simulator sickness so that breaks between navigation tasks were needed. Consequently, the removal of data points from this analysis seemed justified.

To further substantiate our hypotheses of equivalence, an additional analysis was conducted which uses the amount of overlap of inferential confidence intervals (ICIs). Overlap between scores of both groups was very low for all variables. The analysis revealed that a substantial amount of data points were located at the extreme positive and negative ends of the parameters' distributions. The finding of such small overlap of our groups in light of no significant differences and small effect sizes suggests that further research is needed to explain navigation behavior in complex natural environments. Equivalence of navigation in real and VEs would result in high overlap of ICIs. On the contrary, substantial differences between both navigation scenarios would lead to findings of large effect sizes and significant differences between navigation parameters. Our results provide no clear evidence for either scenario, so that an extended evaluation of this paradigm is required to shed light on the relationship of virtual and real navigation behavior. The conducted analyses suggest that a large variability of navigation behavior is evident in complex real-world and virtual scenarios. User interaction and visual properties of the VE (Bowman, Koller, & Hodges, 1998; Stanney, et al., 2002; Stanney, Mollaghasemi, Reeves, Breaux, & Graeber, 2003) play an important role in any comparison of virtual and real navigation. Future studies need to explore which of these factors contribute to such variability in either scenario.

A variety of other measures were used to quantify the participants' ability to find their way through the building and estimate the position of targets around them. None of these measures produced a significant difference, but effect sizes varied considerably between tasks. Distances in the VE were consistently underestimated which is in line with previous findings in the literature (Furness & Henry, 1993; Witmer & Kline, 1998). However, contrary to other experiments, our targets were not visible from the participant's viewpoint and had to be judged based on configurational knowledge of the building rather than visual cues. Similar to the study by Furness and Henry (1993), display distortion could play a role in the participants' underestimations of distance. In their study a group of 24 architects were asked to complete navigation, orientation and distance judgment tasks in a real and virtual gallery under several different display conditions. The authors report that underestimations were greatest when a head-tracked head-mounted display was used which potentially led to the participants seeing a distorted image at the periphery of the projected field-of-view. The projection screens in this current experiment were also distorted towards the left and right horizontal edges which in turn can potentially affect orientation behavior and distance judgment.

The number of stops and the total time participants stopped on their routes were intended to assess the extent to which each person showed orientation behavior. Participants in the real building used such stops to search for landmarks and find their bearings. Unfortunately, many additional virtual stops were recorded due to difficulties with the computer mouse and issues with collision detection within the VE. With these limitations in mind, it comes as no surprise that a significant difference for both variables of stopping behavior was observed and the results of these analyses cannot contribute to the interpretation of navigation ability as intended.

In order to consider the use of VEs in cognitive rehabilitation the difficulty of three-dimensional environments has to be evaluated. Such quantification of difficulty is necessary to provide alternate versions of navigation tasks, classify routes and environments which patients are exposed to, and adjust training difficulty in the context of complex rehabilitation trainings. In recent experiments, researchers manipulated the number of turns or the length of the route, because they are simple to measure and implement in an experimental setting (Koh, et al., 1999; Ruddle, et al., 1997). However, most real-world environments cannot be compared to simple office corridors in a university building. Everyday scenarios like residential houses or shopping malls often have multiple floors and there is more than one viable path which leads to the target. For simulation of these scenarios different measures need to be found in order to assess complex behavior in a standardized, systematic way. How visibility, number of possible routes, results of pathfinding algorithms or other yet undefined variables influence the navigation performance in complex environments must be subject to further investigation. In addition, the relationship of VE performance and well-established clinical measures of spatial abilities needs to be examined. The results of Nadolne and Stringer (2001) and Kozhevnikov and Hegarty (2001) suggest that small-scale tasks like mental rotation place different demands on the cognitive system than navigating through the environment. Hence, more ecologically valid assessments are needed and the continued evaluation of VEs for such purpose seems justified. Navigation through relevant environments is an important aspect of independent living for brain-injured patients. Assessing and training navigation in complex everyday scenarios as part of context-sensitive rehabilitation is a notable improvement over the current practice of time-consuming training at the hospital or the patient's home. Virtual navigation applications allow for more sophisticated, quantifiable scenarios that can be employed remotely without risk for the patient and lower costs for the therapist and health care provider.

In conclusion, the VNT has shown potential as a useful tool for accurately capturing a complex skill like navigation ability. By leveraging the strengths of VEs the capture, interpretation, and visualization of navigation data has been achieved. However, our results show no correlations with other measures of spatial ability and the complexity and high variability of our data did not allow for an unambiguous interpretation. This suggests that measuring navigation ability in all its facets is a highly complex matter which cannot easily be related to existing measures of configurational knowledge of environments. To further increase the validity of gathered navigation data, several improvements towards higher usability of the VE are necessary. Issues of simulator sickness, display distortion and model detail of the VE need to be addressed. Further, it is necessary to replicate the experiment with brain-injured patients in order to understand whether navigation behavior is equivalent in real and virtual environments for this clinical population. Current results cannot be used to draw conclusions about the navigation performance of brain-injured patients. Previous studies suggest that brain-injured patients perform significantly worse on almost all tests of cognitive function, depending on the severity and location of brain lesions (Rao, et al., 1999; Schretlen & Shapiro, 2003). Unfortunately, only little evidence was found that specifically addresses differences in navigation behavior in healthy and brain-injured individuals. Only Livingstone and Skelton (2007) found that patients with traumatic brain injuries were significantly worse than healthy controls in using distant landmarks in the background to navigate through a virtual model of the Morris Water Maze. This suggests that the patients' frontal lobe and hippocampus lesions may impair the ability to form mental maps of the environment. Consequently, brain-injured patients might show significantly worse performance on the pointing task applied in this thesis (i.e. pointing to distal unseen landmarks) and also require more detailed proximal cues in the VE in order to find the target locations. Though it is currently unknown how brain injuries affect patients' performance in a complex environment such as the Erskine building.

However, the current investigation provides a first step and comparison data for future studies with a clinical sample. With refined navigation measures and large samples of brain-injured participants more insights into the underlying factors of navigation performance variability are expected. Such insights are needed to utilize more ecologically valid assessments, as with higher ecological validity complexity of the assessment increases substantially. A valid cognitive task with such high ecological validity could greatly contribute to context-sensitive rehabilitation and give the clinician a basis for judging the patient's abilities in real-life settings.

3.2 Experiment 2 - Evaluation of Virtual Memory Task (VMT)

This study has been conducted at the Neurology Department of the Asklepios Rehabilitation Clinic, Schaufling, Germany, between October 2010 and April 2011. Results of this clinical trial have been published at the VR International Conference 2011, Laval, France (Koenig, Crucian, Dünser, Bartneck, & Dalrymple-Alford, 2011a) and the International Journal of Design and Innovation Research (Koenig, Crucian, Dünser, Bartneck, & Dalrymple-Alford, 2011b). The main goal of this experimental trial initially was the validation of the aforementioned VMT (see Chapter 2.4.3) in order for it to be integrated into context-sensitive rehabilitation. However, the inclusion of patients with a wide range of neurological disorders also provided valuable feedback about using a VE-based cognitive task with such population. At the same time, the heterogeneous group of participants did not allow for firm conclusions about the validity of the employed methods.

3.2.1 Introduction

During inpatient rehabilitation therapists are often faced with the uncertainty of how the patient is going to perform at home or the workplace after rehabilitation ends. Assessments usually reflect patient performance at the clinic in a highly structured environment. This performance may not translate to the home environment or workplace. In their daily routine individuals are faced with decisions, obstacles and unpredictable situations that often exceed complexity of structured therapies and activities in a clinical environment. Unless an outpatient program is planned, the clinical team typically does not receive any feedback about how the patient fares during daily activities. In worst case scenarios, patients return to the clinic after their situation worsened and they have failed to live independently. Until now, an ecologically valid task which reflects the individual circumstances of the patient seemed unrealistic in terms of required labor, construct validity, and cost-efficiency. Such task would give the clinical team a basis for deciding about the patient's aftercare and day-to-day performance.

An additional challenge during cognitive rehabilitation is the patients' motivation for engaging in highly repetitive training tasks. Task repetition in different contexts is especially important to promote generalization of practiced skills (Sohlberg & Mateer, 2001). However, repeating monotonous tasks several times each day or going through abstract batteries of cognitive assessments can adversely impact a patient's motivation. Weak motivation is even more likely if the training tasks are not relevant to the patient's daily life (i.e. low ecological validity). This is often the case with simplified cognitive tasks where attention or memory functions are trained with shapes, patterns or primitives (e.g. several subtasks of CogPack[31] and RehaCom[32]). Modern commercial cognitive tasks, brain teasers, and

---

[31] Marker software – CogPack – www.cogpack.de

[32] Hasomed RehaCom – www.hasomed.de

games are more entertaining and subjectively seem to show higher face validity, but scientific evidence for functional improvement in user groups with brain injuries is sparse (Westerberg et al., 2007). Further evidence is necessary to ensure that these programs support generalization of trained abilities (Owen et al., 2010). Consequently, a set of validated cognitive tasks which are relevant to the patient's needs and background are needed.

This present study proposes a cognitive task which can be integrated into many VEs. The VMT was designed to provide a clinical tool which has several advantages over traditional cognitive tasks:

- higher ecological validity by using personalized, realistic VEs
- higher motivation for patients to practice the task frequently in a meaningful test environment
- precise measurements in three-dimensional space for analyzing the task's results

The proposed VMT has been tested with 45 individuals with a wide range of neuropsychological deficits at the neurological department of a German rehabilitation clinic. It was designed to involve a combination of short-term memory and perspective taking skills. Hence, correlations with neuropsychological tests measuring those constructs are expected. Specifically, the VMT's outcome measure is predicted to significantly correlate with pencil and paper measures of spatial abilities and visual short-term memory. No significant correlation is predicted with measures of attention. Further, it is expected that a larger number of target objects and a larger perspective change are associated with larger VMT error scores. The test's integration into clinical context, usability, task development, and task validation are discussed in the following chapter.

3.2.2 Methods

3.2.2.1 Participants

45 participants (22 male - 23 female) at the Neurological Department of the Asklepios Clinic Schaufling, Germany, were recruited for this trial. Neurological patients with severe traumatic brain injury (6 patients), subarachnoidal hemorrhage (2), brain tumor (4), epilepsy (5, including 2 with hippocampal sclerosis), stroke (9, mostly right-hemispheric), normal pressure hydrocephalus (1), Chorea Huntington (1), Syringomyelia (1), Multiple Sclerosis (6), anaphylactic shock (1), herpes encephalitis (1), meningitis (1), and hypoxic brain damage (1) volunteered to participate in this study. Volunteers were specifically chosen to represent a broad range of attentional and mnestic deficits, including non-deficient and highly-impaired individuals. Five therapists and one orthopedic patient without cognitive deficits were also recruited for this study. Average age of the participants was 38.56 years (range 17 – 66 years). The only requirement for

recruitment was the ability to concentrate and maintain performance for at least 30 minutes. Computer experience was not required for participation. All patients were able to give informed written consent. Availability of patients, range and severity of brain injuries and inclusion and exclusion criteria were discussed with the clinical staff at the hospital. Due to the limited timeframe of six months, the anticipated availability of patients and the unpredictable nature of neurological rehabilitation (dropout due to return to acute hospital, occurrence of seizures, headaches, refusal for informed consent, lack of health plan coverage) patient numbers cannot reliably be estimated. Hence, specific brain lesions (e.g. right temporo-parietal lesions for deficient spatial abilities) were not taken into account as inclusion/exclusion criteria in order to avoid receiving not enough feedback and only few data points for statistical and usability evaluation.

3.2.2.2 Design

Order of tasks was identical for most patients but differed in few cases when patients had already completed tests with other therapists or at previous hospitals. Assessments were completed within a few days up to three weeks, depending on the patient's therapy schedule.

3.2.2.3 Materials

3.2.2.3.1 Pen and Paper Tests

All pen and paper tests were translated to German where appropriate. For economic reasons translations were done by the experimenter and discussed with the clinical team at the Asklepios Clinic, Schaufling, for their appropriateness and use of suitable terminology.

Spatial abilities were assessed using the Object Perspective Taking Test (Hegarty & Waller, 2004) and the Mental Rotations Test (Vandenberg & Kuse, 1978). Attention was assessed with the D2 Test of Attention (Brickenkamp, 1981). Memory and working memory assessment consisted of the immediate forward and backward block span and digit span of the Wechsler Memory Scale III (Wechsler, 1945/1997), and the Rey-Osterrieth Complex Figure Test (Osterrieth, 1944). An adapted version of the Computer/Internet Experience and Skills Questionnaire for: Internet Diabetes Trial at Harborview (Goldberg, 2006) was used to assess computer experience and skills. German versions and translations were used for all test instruments.

The OPTT and MRT have already been described in Experiment 1. Details about both tests can be found in the previous chapter.

Block and digit span assess the visual and verbal short-term memory (forward) and working memory (backward). The experimenter taps on a sequence of blocks or reads a sequence of digits which the participant has to reproduce in the same or reversed order (forward/backward). Difficulty is increased gradually across trials. Correctly reproduced items provide the test score.

The Rey-Osterrieth Complex Figure Test consists of a complex drawing which can be decomposed into 18 distinct objects. The participant's first task is to copy the reference figure without omitting any details. In a second trial, the participant has to immediately draw the figure from memory once the reference drawing has been removed from sight. A third trial has to be completed after 30 minutes in which the participant has to once more draw the figure from memory. Immediate and delayed recall trials were scored and analyzed separately, because due to time restraints not all participants were able to complete the delayed trial.

The D2 Test of Attention consists of 14 rows of stimuli on a DIN A4 sheet of paper, each consisting of 47 letters ("d" or "p"). Additionally, each letter is accompanied by a series of dashes above or below the letter. The participant's task is to identify each target "d" containing a total of two dashes, either above, below the letter or both. The participant is given 20 seconds per row to identify as many of the 21 or 22 targets as possible. Stimuli are processed consecutively within each row. After each 20 second interval, the experimenter gives a cue to advance to the beginning of the next row. Results are analyzed for processing speed, omission and false positive errors. The total number of processed targets minus the number of errors is used as the test's score.

3.2.2.3.2 VMT

The VMT was placed in a realistic, to scale model of the rehabilitation clinic in which the study was carried out. Most rooms of the clinic have been modeled for carrying out several experiments. Only one virtual office room within the clinic was chosen for the VMT assessment. Sufficient detail and photorealistic textures were used in order to enable participants to easily recognize the environment. The 3D model was created using Google SketchUp 8 Pro. Textures were imported from photographs, prepared with Genetica 3.51 Basic Edition and used within Google SketchUp. Measurements for accurate modeling were gathered manually from the real environment. Interactivity of the environment, data collection and task logic were implemented using the game engine Unity, Pro version 3.1. Task development and testing procedures were carried out on a PC workstation with AMD hexacore CPU, 2GB NVIDIA GTX460 graphics card, 8GB of memory and solid state drive. All tasks were displayed on a 24-inch LCD monitor that was placed 60cm in front of the participant. Keyboard and mouse were used to interact with all tasks. Development of the VE followed the procedures as outlined in previous chapters (see Chapter 2.3).

The VMT was implemented in the virtual model of the office in which the participant was seated during all tests. Real and virtual viewpoints were identical so that the participant was facing the same 90cm x 100cm virtual table on which keyboard, mouse and monitor were placed. The virtual table was empty apart from several task-relevant items (Figure 14). The virtual office room was deliberately chosen for this experiment, because it was easily accessible for 3D modeling. Further, it was one of few rooms in the clinic to which patients did not have any previous exposure before the test session was conducted. For the purpose of this study the replication of the office room was not strictly necessary, as the validity evaluation could have taken place in any VE. However, the virtual office did give the participants an additional sense of space. Moreover, the use of a generic VE was avoided, because the virtual model of the rehabilitation clinic (including the office room) was also needed for experiment three (chapter 3.3).

Prior to the first task the participant was shown an overview of the surrounding VE for 15 seconds to allow for better orientation within the virtual office. Instructions were given to focus attention on the virtual table and the items placed on the table. The participant was given two minutes to memorize the exact locations of the target items. After two minutes or as soon as the participant indicated that all locations had been memorized, the target objects were moved to new locations on the table. Locations for all trials were initially randomized during test development and identical for all participants. The participant's task was to precisely drag and drop the items back to the initially learned locations. Each trial included a specific number of target items (4, 5, 6 or 7) and a defined change in perspective. The initial perspective while learning the item locations was always congruent with the participant's viewpoint (Figure 15). When items were moved to new locations, the perspective either remained unchanged, moved to the left of the table (90 degree shift) or to the opposite side of the table (180 degree shift). The viewpoint change was carried out as a passive, continuous motion towards a new location in the VE with the user's virtual field of view always centered on the virtual table. The participant had no control over the viewpoint change at any given time.

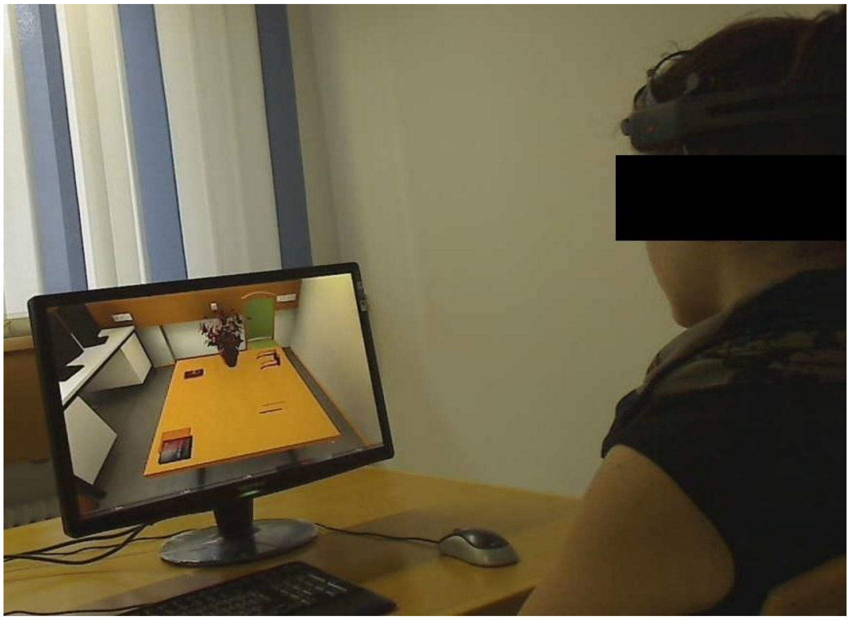
[redacted]

Figure 14. Participant completing the VMT

Participants were not informed about upcoming perspective changes and were instructed to take into account possible perspective changes when learning the spatial layout of items. Even though the participants were allowed to look around within the VE, the viewpoint could not be changed far enough to give any cues about the original perspective before the items were moved. Target items consisted of two sets of objects which alternated between trials and included typical items in an office environment (e.g. book, cup, bottle, trash can, pencil).

Number of target items increased gradually from four to seven. Each participant went through the same order of twelve trials which were a combination of three perspectives (0, 90 & 180 degrees) for each of the four numbers of items (4, 5, 6 & 7 items). Target items were selected and moved using the left mouse button by dragging the object to a new position. The experimenter used a keyboard to manually select items when the participant had problems using the mouse. This was evident for almost all participants when very small items had to be selected (e.g. pencil).

A distance error score was calculated for each target by finding the distance between the participant's answer and the item's original position during the learning phase, measured in meters. The largest possible error score on the virtual table was approximately one meter. Distance error and all target positions were saved as text files for each trial. Rotation of target items was not relevant for this experiment.

Prior to this study, several brain-injured patients and therapists took part in preliminary usability trials to test the application and a variety of user interfaces. A combination of mouse and keyboard controls emerged as the preferred alternative. 44 of the 45 tested participants were able to effortlessly control the application and drag the target items to their original locations without any instructions at all. Even patients with little computer experience were easily capable of selecting and moving items. The experimenter used the keyboard to manually select target items whenever participants struggled to click on smaller objects. In cases of severe cognitive impairments or aversion of computer technology, the experimenter is also able to move the targets via instructions by the participant. This was done for one patient with no computer experience at all. Patients with severe motor deficits can use a modified USB-numeric keypad with large keys to move targets onscreen. The keypad was initially planned as a backup input device, but was not used in any of the described trials.

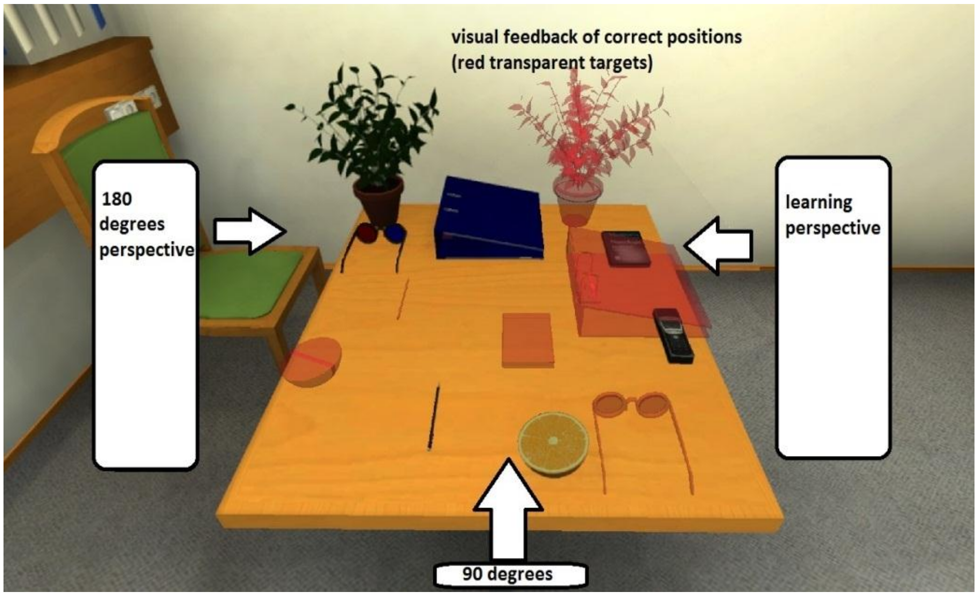

Figure 15. Perspective changes for Virtual Memory Task

3.2.2.4 Procedure

Patients were identified during admission at the clinic and approached during an initial meeting with the clinical team. In a second 30-minute session the study was explained in detail and informed written consent was established. Most patients completed the assessments in 120 to 180 minutes spread across three to four sessions, each lasting 30 or 60 minutes, depending on the patient's schedule and constitution. Sessions were carried out in addition to the normal therapeutic schedule of the patient. During a first one-hour session (or two 30-minute sessions), computer experience, block and digit span of the WMS III, Object Perspective Taking Test, and Rey-Osterrieth Complex Figure Test were completed. In addition, a target pointing and several orientation tasks were used to assess knowledge of the clinic buildings. Analysis of the orientation tasks (mental maps task) will be discussed as part of Experiment 3. In subsequent sessions all participants were assessed with D2 Test of Attention, Mental Rotations Test and the computerized VMT. All individual results were immediately analyzed and feedback was given to the patient after each session. Not all patients were able to complete the full experimental protocol due to time restrictions or patients being transferred or discharged from the rehabilitation clinic.

3.2.2.5 Statistical Analysis

Statistical analyses were carried out using the software PASW 18[33]. Initial analyses of VMT-results revealed that the assumption of normality of underlying populations has been violated for all test results. QQ-Plots and significant results for Shapiro-Wilk-Tests clearly indicated the non-normal distribution of the population from which our data was drawn. Using Levene's Tests, homogeneous variances of our data sets could only be found after test results from several highly impaired participants were removed from the analyses. A total of three patients showed very large variability in their responses and hence were identified for possible removal of their results from the dataset. However, given the exploratory nature of this study, it was decided to not remove any data and use non-parametric tests instead. Consequently, performance on the VMT was analyzed using the non-parametric Friedman test for repeated measures analysis. Both of the VMT's factors were collapsed and analyzed individually. Interactions of both factors have not been addressed in this study. For post-hoc analyses, Bonferroni-adjusted Wilcoxon Signed-Rank tests were used to find differences between each test condition. Spearman rank-order correlation coefficients were calculated for the results of the VMT and all other cognitive tests to assess discriminant and convergent validity. Bonferroni corrections were used to adjust the α-level for multiple comparisons.

3.2.3 Results

The participants' performance on the VMT was subject to two Friedman tests, analyzing each of the test's two factors separately – perspective change (Figure 16)

---

[33] IBM SPSS Statistics – http://www-01.ibm.com/software/analytics/spss/

and number of target items (Figure 17). When comparing trials with different numbers of target items, a significant difference in memory performance became apparent (λ (3) = 27.32, $p < 0.001$). Trials with different perspective changes also differed significantly ($\lambda(2) = 42.19$, $p < 0.001$).

Given that both test factors were expected to increase complexity of the testing situation, it was hypothesized that distance error, which was dependent variable of all VMT-analyses, would also increase gradually as number of target items and angular perspective change increase. Consequently, one-sided Wilcoxon Signed-Rank tests were used to compare all individual conditions for each factor. Bonferroni adjustments of the α-level were employed for all tests (α = 5%/9 = 0.55%). Significant differences were found for comparisons of number of target items between five and seven targets ($z = -3.48$, $p < 0.001$, $n_5=43$, $n_7=37$, difference in mean rank = 2.11) and six and seven targets ($z = -4.02$, $p = 0.001$, $n_6=40$, $n_7=37$, difference in mean rank = 2.31). Memory performance in respect to perspective changes differed significantly between 0 and 90 degrees ($z = -5.10$, $p < 0.001$, $n_0=n_{90}=43$, difference in mean rank = 15.69) and between 0 and 180 degrees ($z = -5.43$, $p < 0.001$, $n_0=n_{180}=43$, difference in mean rank = 17.92). All remaining pairwise comparisons did not show significant differences in the predicted direction.

Total distance errors across all perspective changes and number of targets were correlated with results from Rey-Osterrieth Complex Figure Test, digit and blockspan, computer experience and D2 Test of Attention. Average distance errors across all trials with a changed perspective were correlated with Mental Rotations Test and Object Perspective Taking Test. Spearman's rank-order correlation coefficient was used with an adjusted α-level of 0.5% (α = 5%/10). Strong significant relationships were found between the VMT-scores and immediate and delayed recall of the Rey-Osterrieth Complex Figure Test (one-sided test). No significant relationships were found for VMT results and computer experience or VMT and D2 Test of Attention. Only the latter two correlations were analyzed with two-tailed tests, because there is no rationale for either a positive or a negative relationship and no significant correlations were expected. Tests of spatial abilities were not significantly correlated with VMT scores. Detailed results can be found in Table 5.

Table 5. Spearman's rank-order correlations

Test	**N**	**Spearman's rho**	**p**
Rey-O. Complex Figure immediate recall	21	-0.76	<0.001[a]
Rey-O. Complex Figure delayed recall	19	-0.76	<0.001[a]
DigitSpan Forward	36	-0.36	0.032
DigitSpan Backward	36	-0.36	0.030
BlockSpan Forward	36	-0.20	0.247
BlockSpan Backward	36	-0,26	0.119
Computer Experience	32	-0.20	0.272
D2 Test of Attention	22	-0.39	0.072
Object Perspective Taking Test	38	0.40	0.013
Mental Rotations Test	18	-0.38	0.122

a – indicates a significant correlation at $p<0.005$ (p = 0.05/10)

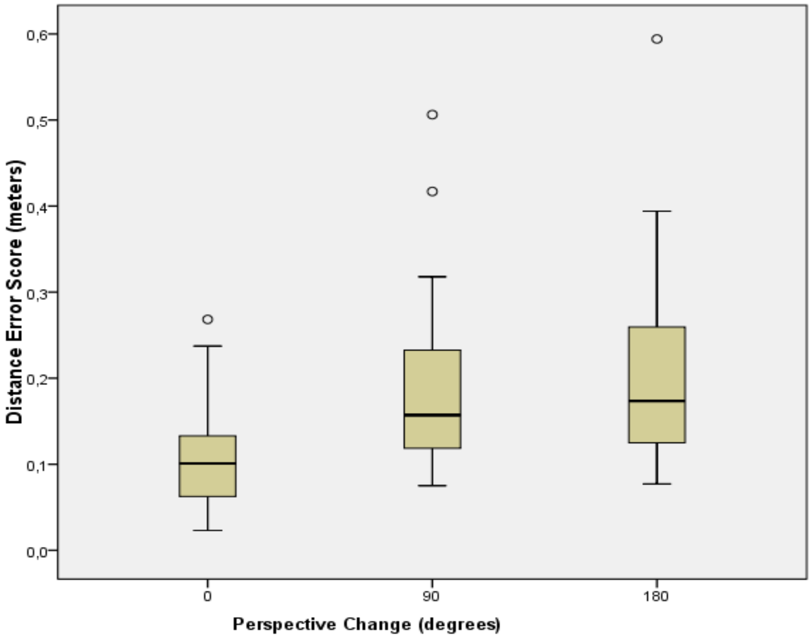

Figure 16. Boxplot of Distance Error Score and Perspective Change

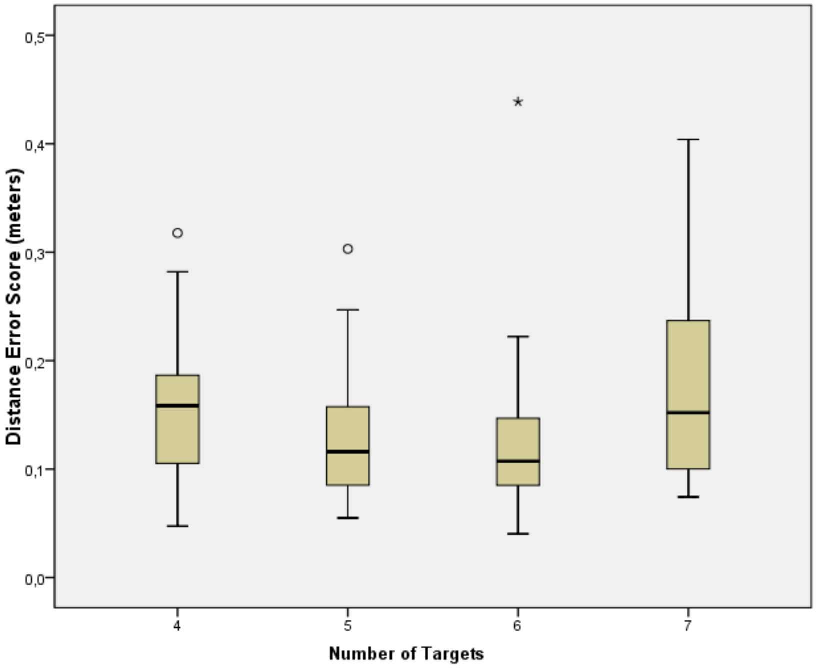

Figure 17. Boxplot of Distance Error Score and Number of Targets

3.2.4 Discussion

The aim of this present study was to demonstrate the viability of a modular cognitive task which was implemented in a virtual model of a rehabilitation clinic. Primarily, this experiment intended to demonstrate the test's convergent and discriminant validity in a clinical context and to receive usability feedback from a wide range of brain-injured patients. The task assessed short-term memory and the ability to imagine different perspectives in three-dimensional space. 45 patients with a broad range of cognitive deficits were included in this trial to compare test results with established neuropsychological tests. Further, the participants' feedback was used to improve interface usability. The VMT was designed to allow clinicians and researchers to individually target deficits and context of each patient. The test was created to provide a clinical tool with higher ecological validity than existing tasks. Further, the test's setting in three-dimensional virtual space allows for exact measurements and differentiated visual and statistical analysis of test results. As such, the test can be integrated into any VE that can be imported and displayed within the game engine Unity. For the purpose of this experiment the VMT was implemented in a virtual model of the experimenter's office. The location was chosen to replicate the real room where the actual experiment was conducted.

It was expected that test performance on the VMT shows strong significant correlations with measures of short-term memory (Hypothesis I-b) and visual memory (Hypothesis I-d). It was also expected that for trials during which a perspective change is applied, a significant correlation with tests of spatial abilities is evident. No significant correlation was expected between VMT performance and measures of attention (Hypothesis I-e). This trial did not assess the extent to which the VMT possesses ecological validity. Further, it was not evaluated to which extent the test can be adjusted to fit the individual goals and deficits of each patient. Rather, the study was a standardized protocol to assess the test's validity and feasibility. Ecological relevance and the test's flexibility were evaluated in experiments three and four. Based on previous trials conducted by Kozhevnikov and Hegarty (2001) and King and colleagues (2002) the overall difficulty of the combination of memory and spatial abilities could not be anticipated. However, based on the task's relevance for everyday life, it was expected that both task factors combined provide an appropriate level of difficulty. This expectation was based on the assumption that everyday objects and relevant targets often are dynamic and viewed from multiple perspectives. For example, walking through a novel environment multiple times will presumably result in different routes, relevant locations and perspectives each time – exemplarily in the case of searching for a parked car in a complex scene.

As hypothesized, the VMT error scores showed a strong negative correlation with scores of the Rey-Osterrieth-Complex-Figure Test (immediate and delayed recall). Both tests make high demands on visual memory, so that convergent validity has been established for the VMT and Hypothesis I-c is confirmed. No significant relationship has been shown between VMT error scores and tests of spatial abilities (i.e. Hypothesis I-d is not supported). Even after excluding trials without viewpoint changes and each participant's first trial with a viewpoint change from the analysis (as most participants were surprised by the rotation), no significant correlations were obtained. However, variability for test results of the Mental Rotations Test and Object Perspective Taking Test has been very high. Due to the small sample size, N=18 and N=38 for Mental Rotations Test and Object Perspective Taking Test respectively, and the heterogeneous sample of neurological patients, further investigations are necessary to establish possible relationships with tests of spatial abilities. No significant correlations were found for VMT results and digit or block span. While block span and VMT both are expected to assess the construct of visual working memory, the concept of test scores differs between both tests. Digit and block span count each test item as either correct or wrong. Results of the VMT provide much more information so that exact positions of each target item can be calculated in 3D space. This allows for differentiated analyses for several types of errors. Errors can occur for rotation of the array of items (Figure 18-A; i.e. ignoring a perspective change), distance between targets (Figure 18-B; with correct layout of targets), total shift of the array of items (Figure 18-C; e.g. when misinterpreting foreshortening of camera perspective), swapping target locations (Figure 18-D), or location of single targets (Figure 18-E; *"I forgot where it was"*).

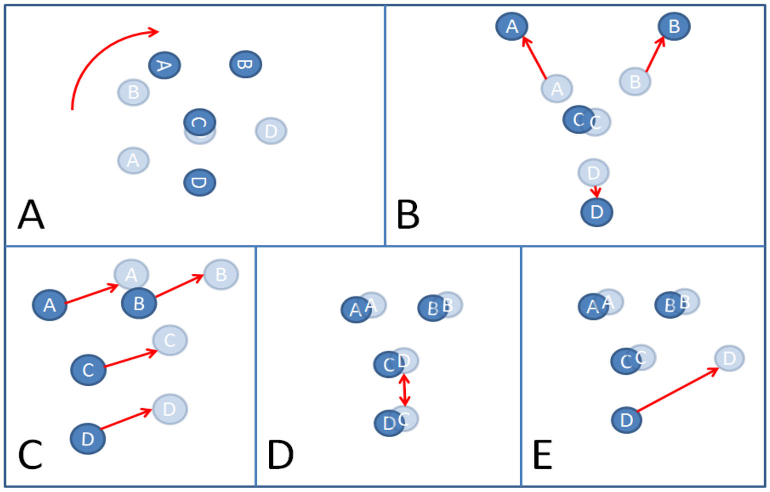

Figure 18. Observed Error Types for VMT

After experiment two was finished, several prototypical applications for data analysis were developed. During the development process of these analysis tools, errors D and E (Figure 18) emerged as the most common error types. Combination of error types did also occur frequently. However, unless an analysis tool has been finished or a simple dichotomy of correct or false answers has been found for the VMT results, no direct comparison to digit and block span seems possible. For the purpose of this study, the absolute distance of the user's answer to the correct (changed) position of the target was measured. Several alternative approaches to error analyses were tested, but none provided satisfying sensitivity for error types. For example, differences between user answers and correct positions were calculated as deviations from the common midpoint of the set of targets on a two-dimensional grid. Figure 19 shows an example of an analysis tool which calculates the proportional difference in distances between original target location and correct answer and the user's answer respectively. The application has been specifically developed to run in the Unity editor and read the text files that the VMT produces. It is intended to integrate this application into the standalone executable of the VMT so that clinicians can visualize the task results without the need to run the Unity editor. The tool can visualize the positions of the user's answers and the correct target locations. Future trials will be directed at making more extensive use of the large amount of data that the proposed virtual task produces.

Degrees: 90.00002

Config Error: 0.932398

Figure 19. Unity-based analysis tool for the VMT

Types of errors also appeared to be related to the participants' strategy for memorizing the target locations. Participants reported strategies about using marks on the wooden (virtual) table, using external cues (e.g. positions of power outlets and chairs in the scene's background), learning the relative positions of targets to each other or simply using a mental picture of the whole scene. Unfortunately, strategy use was not recorded for each participant so that a relationship between both variables could not be established. Such additional information is expected to broaden the use of the proposed cognitive task by enabling the therapist to teach new strategies to patients after they suffered from brain injury.

Correlation analysis of test scores for the D2 Test of Attention and VMT did not reveal any significant relationship. While several patients did show severe attention deficits, no linear relationship was expected between both tests. In order to confirm discriminant validity of the VMT, results of this study need to be replicated with larger and more homogeneous samples of neurological patients and healthy adults.

The heterogeneous sample was specifically chosen to represent patients with a wide range of cognitive deficits and purposely included both, healthy and impaired individuals. While strict test validation is an important aspect which needs to be expanded upon in upcoming trials, an important goal of this present study was to explore the usage of such virtual task in a clinical context. Hence, a more stringent selection of participants in future studies (e.g. right-hemispheric temporal/parietal lesions, healthy control group) is necessary to draw conclusions about construct validity and relationship to other psychometric measures. Further, the memory performance of healthy and brain-injured individuals in more homogeneous samples could shed light on the controversial role of the human hippocampus in spatial memory. Either general memory load, as suggested by

Shrager and colleagues (2006), or allocentric viewpoint changes (King, et al., 2002), have been associated with the human hippocampus. The VMT builds upon the tasks of both groups and extends them for use in everyday clinical training and assessment. The VMT's results suggest that memory performance is influenced by viewpoint change and memory load (i.e. number of target items). However, the interaction of both factors needs to be evaluated in future trials with less heterogeneous samples.

Several other task-related aspects need to be evaluated in future trials. To address a possible confounding factor of incongruent perspectives between real and VEs, it is necessary to use separate locations for VE and physical space. Changing the virtual perspective caused confusion among several participants whenever the virtual rotated perspective conflicted with the real perspective of the participant sitting in front of the table. The match between real and virtual test environment was supposed to give the participant a better sense of space during the experiment. Also, the office room was one of few rooms to which the experimenter had easy access for measurements and 3D modeling. Further, within the office setting patients' exposure to the actual environment was easily controlled by the experimenter. All participants were exposed to the office for approximately 60 minutes prior to the use of the VMT (30-minute information session, 30-minute testing session). Future studies will have to address this confounding variable and evaluate or control the interference that such overlap of real and VEs may cause.

The choice of target items for the VMT is an important parameter which needs to be controlled in future studies. The task is designed to allow the clinician to choose targets which are of personal relevance to the patient. However, it is unclear whether the familiarity and repeated use of target items in consecutive trials have an effect on the task's results. For the purpose of this study, item sets were always alternated between trials. Thus, the occurrence of false memories (i.e. item positions from past trials) is a possible confounding factor for this study's results.

An additional goal of this study was to show that both of the task's factors, perspective change and number of target items, contribute towards the difficulty of each trial. The nonparametric analyses revealed that participants committed the largest errors for trials with seven target items and generally more errors for trials with larger perspective changes. No significant differences were found between trials with five and six target items and between trials with four target items and all other trials. Even though the task was explained in detail to each participant prior to starting the first four-item-trials, several participants were surprised by the task's mechanics, especially perspective changes. It can be assumed that task performance on the first trials did not reflect the participant's true abilities, but rather was affected by the novelty of the task. A practice trial or an initial simulation of perspective changes need to be considered for future trials. The absence of a significant difference between trials with five and six target items can possibly be explained by the use of biased item constellations. Item locations were randomly chosen during the VMT's development cycle by using a random number generator and thus could falsely result in non-linear difficulty progression throughout the testing process. As a consequence, the influence of target distances and complex item constellations (e.g. items being easier to remember by standing close to each other or in a triangle arrangement) on trial difficulty need to be evaluated in future trials.

The current study provided a first-hand experience of how participants act in a realistic, semi-familiar VE. Further, the office setting enabled the experimenter to draw comparisons to the real environment whenever participants were skeptical about test results or the nature of the task. These situations were crucial for showing the effects of a cognitive task with high ecological validity. While this comparison cannot be made during day-to-day clinical use of the test, it was beneficial for clearly showing the task's transparency and evoking patient responses that support the test's ecological relevance. More specifically, this trial provides evidence for the strong effect that the VMT's high ecological relevance had on the patients' awareness of their cognitive deficits. Five cases were identified in which patients with mnestic deficits went through all well-established pencil and paper assessments with constant denial of their deficits. Even when faced with extremely poor test results on the Rey-Osterrieth-Complex Figure Test or block and digit span, no deficit awareness was evident. However, when these participants were assessed with the VMT, they were confronted with a task that is believable, transparent and easily comparable to relevant tasks of daily life. When faced with their poor results on the VMT, emotional outbursts and breakdowns were evoked, mostly among five out of 45 participants. To illustrate this further, one of the participants concluded, *"I can't believe I'm not able to do this. Even a [expletive] third-grader can do this".* However, it is important to note that these reactions require additional care when administering such virtual task. Awareness of deficits is a vital aspect of cognitive rehabilitation, but without proper support from an experienced therapist during and after an emotional experience like this, the positive outcome of a patient's rehabilitation is at stake. This also leads to the conclusion that unguided use of such virtual task without the patient's insight into their own deficits is not recommended at this stage and therapeutic potential of tasks with high ecological relevance needs to be extensively tested in future trials. Though, from a clinical perspective it is to be expected that patients who are still denying their cognitive deficits when being discharged from a rehabilitation clinic will struggle when returning to the challenges of daily life. Hence, an early confrontation with their own deficits appears to be in the patient's best interest in order to actively partake in rehabilitation. It is such confrontations in a controlled environment (i.e. under supervision of a clinician) which can make a realistic VR test a tremendously helpful tool during and after inpatient rehabilitation. Hence, patients could be introduced to the VMT during inpatient rehabilitation and complete multiple sessions during which they solve tasks of varying difficulty. Once the patient is familiar with the test and it can be assumed that the patient is aware of his/her own deficits based on verbal assessment, unsupervised use within the clinic and outside of the clinic could be considered over time. Future trials need to more systematically explore the effect of developed cognitive tests on deficit insights. More specifically, test sessions need to be recorded to provide evidence which task components or test results can influence a patient's deficit denial. The presented anecdotal evidence can merely serve as first indication of the potential of the VMT and needs further substantiation in future trials.

The development of the VMT task has been focused on creating a modular task which can be easily placed into any VE. The workflow of creating a VE has been refined to allow for quick prototyping of virtual spaces in a matter of hours (see chapter 2.3). The interactive environments are not economical from a performance standpoint so that a high-end PC is currently required to run the applications. However, time is of the essence when a brain-injured patient starts a four to six-week rehabilitation program. The creation of a detailed version of the virtual office in which the assessment has been carried out took a total of four hours. Consequently, it is easily possible to create to scale models of patients' home environments in a matter of one to two days. The exact workflow, performance issues of such VEs and their usage during rehabilitation are described in more detail in the following chapter.

Besides the obvious use of testing and training spatial memory, other application fields for the VMT need to be evaluated in future trials. Target items, item scale and environments can be easily changed to fit the patient's needs. The task-scale can be adjusted to move around furniture or any virtual item. It is also possible to use the application to train memory strategies by repeatedly requiring the patient to place targets at strategic places in the environment. While the targets were always moved to fixed locations after two minutes, learning duration and changed positions can be manipulated by the experimenter. Several parameters have been implemented to randomly move around targets for each trial in order to promote long-term use of the application, e.g. as a training application instead of a diagnostic tool. The transfer of trained skills to tasks of daily life will be evaluated in upcoming clinical trials. Stereoscopic rendering of the VE (anaglyph red/cyan) has been implemented, but was not used for the experiment to avoid unnecessary risk for patients with epilepsy, and eyestrain for patients with nystagmus or other visual deficits. An advantage of the Unity game engine is the uncomplicated use of the application for online assessment. The virtual task and environments can easily be embedded in any html-page. The only aspect of the task which needs to be modified is the process of saving the task results to an online SQL database. This makes it possible for patients to easily continue training after they are discharged from any rehabilitation program and deficit awareness and emotional stability have been achieved.

In conclusion, the proposed VMT has been shown to help several highly-impaired individuals to realize their cognitive deficits. This can be seen as a first indication of the test's ecological validity. Usability and user feedback have been excellent throughout so that further trials and extended use of the application during context-sensitive cognitive rehabilitation seem justified. However, when using the VMT with patients with cognitive impairments, continuous support by experienced therapists is recommended to avoid frustration. The test's transparent nature and realism of the VE appear very helpful for motivating patients, but can also have adverse effects when individuals abruptly realize that their cognitive abilities have suffered during a life-changing neurological event.

3.3 Experiment 3 - Evaluation of workflow and pointing task

VEs (VEs) have seen increasing use in rehabilitation and therapy over the last two decades (Rose, Brooks, & Rizzo, 2005). By combining VEs with immersive displays and natural interfaces, VR applications have become possible for a wide range of therapeutic scenarios. A more detailed overview over VR applications, their clinical usage, advantages and disadvantages has been made in a previous chapter (see chapter 1.5). Despite their advantages, VEs cannot be considered mainstream applications for therapy and rehabilitation yet. Development of VEs is considered a costly and time-consuming process and cost-benefit analyses for VR therapy tools are still rare. The following evaluation has been conducted at the Neurology Department of the Asklepios Rehabilitation Clinic, Schaufling, Germany, between October 2010 and April 2011. The results of this study have been published at the International Conference on Virtual Rehabilitation 2010, Zurich, Switzerland (Koenig, Dünser, Bartneck, Dalrymple-Alford, & Crucian, 2011).

The aim of this experiment is twofold. Firstly, the proposed workflow for creating VEs is tested with four randomly selected rooms. While the workflow is based on a specific set of software which was selected to match the experimenter's experience, this study is aiming to reveal whether this development process is suitable to create high-fidelity VEs within a timeframe that is appropriate for clinical use. The realism of the VEs is then assessed by a recognition survey among staff members of a rehabilitation hospital. It is expected that the realism of the VE is critical for the patient to recognize the environment. After a neurological incident patients often spend several weeks or months in hospitals and rehabilitation clinics. Enabling a patient to train in a familiar environment which they miss and desperately wish to return to is expected to be a very motivating factor.

Secondly, a pointing task has been created to test the spatial orientation ability of the user. It is a goal of this study to evaluate whether pointing task performance is equivalent in a real and VE (Hypothesis I-f). Before the task can be integrated in individualized VEs during cognitive rehabilitation, its ability to assess relevant behavior needs to be tested. Once this equivalence has been established, the test's utility for clinical decision-making can be explored in a larger randomized controlled trial. For experiment three it is specifically hypothesized that each of the four selected VEs can be created in less than a workday (i.e. eight hours; Hypothesis IV-a). Further, each VE is recognized with above chance probability (>50%). For this purpose it is assumed that there is a 50% chance of guessing whether the location was one of the previously shown locations (i.e. forced yes/no choice). Lastly, it is hypothesized that distance, height and angular judgments of the pointing task are equivalent in real and VEs. Despite evidence of underestimating distances in VEs (Witmer & Kline, 1998) the underlying factors of this phenomenon have not been fully explored yet. Due to the high visual quality of the developed VEs, the research hypothesis of equivalence between real and VEs was used.

3.3.1 Workflow Test

To test the proposed workflow four separate locations within the Asklepios Rehabilitation Clinic, Schaufling, Germany, were randomly chosen and modeled. The exact workflow has been described in chapter 2.3. The choice of environments was based on a random selection from all rooms and corridors within the clinic. Each VE and the corresponding real-world photograph can be found in Figure 20.

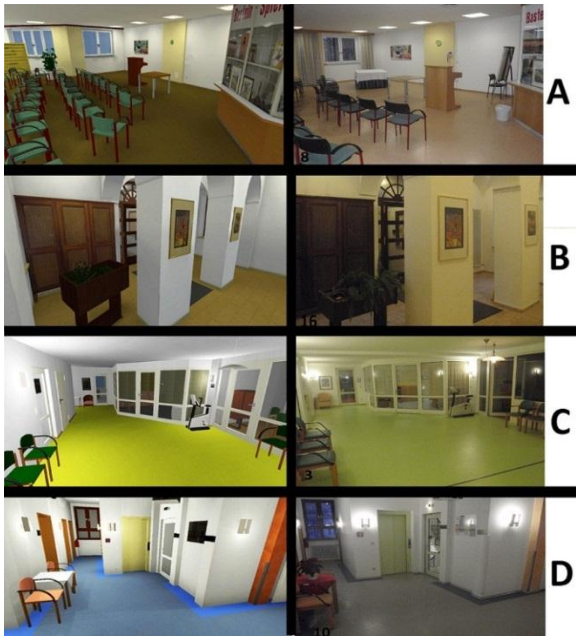

Figure 20. Screenshots (left) and photos (right)

Time to complete each VE was measured from the first photograph to the execution of the finished Unity application (Table 6). All environments were ready for immediate use by a patient. When using this workflow for a patient's home or workplace, the number of rooms, necessary detail and additional travel time to the real location have to be taken into account. Patient and caregivers have to be briefed and the clinical team needs to agree on therapeutic goals in order to implement those goals into the VE. Nonetheless, modeling a small apartment or one floor of a residential house is still possible within the given timeframe of 48 hours after the patient has been admitted to the rehabilitation clinic. With constantly evolving software and a growing library of reusable furniture models, the VEs' realism and pace at which they are created are steadily increasing. However, for the given time estimates it has to be taken into account that the measurements, modeling, and task integration have been conducted by the experimenter who has approximately 2 years of experience with the workflow. If the same development effort was to be done by somebody unfamiliar with the software or proficient with related software, the time to complete each functional environment might vary considerably.

3.3.2 VE Recognition

The realism of all produced VEs has been evaluated by testing how participants recognized each location and how pictures of real and virtual locations were associated with each other. 43 therapists, employees and interns (12 male - 31 female; average age 31.95 years) were shown the four screenshots as depicted in Fig. 19. On average, participants had been working at the clinic for 74.2 months (STD = 91.6, range: 2 weeks - 23 years). Staff members were chosen over brain-injured patients due to the fact that familiar environments are relevant to this thesis. Ultimately, patients are expected to recognize the home or work environment which has been familiar to them for months or years. On the contrary, patients only spend days or few weeks at the rehabilitation hospital. Hence, staff members seemed more appropriate to evaluate the recognition of a familiar environment among distracting items. However, it has to be noted that recognition rates of brain-injured patients cannot be estimated based on this data. Though, anecdotal evidence suggests that a clinical sample similarly recognizes familiar environments unless severe memory deficits prevent the patient from encoding or retrieving the information about the environment. Specifically, all participants except one severely impaired patient recognized the familiar virtual model of the rehabilitation clinic while using the point task (see chapter 3.3.3 below).

Screenshots of the environments were presented on the 7-inch-screen of a Samsung Galaxy Tab tablet PC without time restrictions. Locations A and B were correctly identified by almost all participants (Table 6, identification rate). Locations C and D were more generic and more often mistaken for similar-looking wards and corridors. Because some of the locations were not familiar to interns and new employees, an additional recognition task was used. After showing the four screenshots, 20 real photographs were presented to each employee. Participants were instructed to recognize each scene which was shown on the previous four screenshots. 16 distractor items and the four correct photos (as seen in Figure 20) were shown in the same order for each participant. Recognition rate across all locations was 94.1%, so that almost all participants recognized all four locations (Table 6, recognition rate). During rehabilitation, patients will be faced with very familiar interactive environments and sounds instead of photographs which should further increase the familiarity of the VEs.

Table 6. Results of Recognition Survey

Location (as seen in Fig. 19)	Results			
	Time for Measuring (minutes)	Time for Modeling (minutes)	Identification Rate (percent)	Recognition Rate (percent)
Meeting Room (2-A)	40	162	97.6	97.6
Entrance (2-B)	52	266	90.6	100
Neurology Ward (2-C)	65	270	55.8	93.0
Corridor (2-D)	~50	~250	76.7	88.3

3.3.3 Validity evaluation – Pointing task

As previously described, four VEs were created to demonstrate the effectiveness of the proposed workflow. Before these VEs can be used in a clinical context, cognitive tasks have to be embedded in each environment. To explore such clinical use and investigate the validity of collected data, one of the four environments was chosen for a virtual orientation task. Neurological patients were recruited to interact with the VE and point towards unseen targets in a real and virtual scenario within the clinic building. This test intended to assess the patients' mental representations of their surroundings. If mental maps of real and VEs are equivalent, as assessed through the pointing task, the virtual task's data can easily be used with a patient's individual VE to assess orientation performance in locations outside of the rehabilitation clinic (i.e. home or workplace).

3.3.3.1 Participants

31 patients (16 male, 15 female; average age: 43.9 years; age range: 18-65 years) at the rehabilitation clinic Asklepios Clinic Schaufling, Germany, took part in this clinical trial. Patients with severe traumatic brain injury (3), subarachnoid hemorrhage (2), brain tumor (4), hippocampal sclerosis (1), stroke (8, mostly right-hemispheric), normal pressure hydrocephalus (1), syringomyelia (1), multiple sclerosis (3), anaphylactic shock (1), idiopathic epilepsy (2), cerebral encephalitis (1), cerebral hemorrhage (1) and four patients from non-neurologic departments were recruited for this study. All patients were able to give informed, written consent. Availability of patients, inclusion and exclusion criteria were discussed with clinicians at the hospital. Based on these discussions it was decided to include a wide range of neurological patients instead of limiting the recruitment to patients with specific lesions of the righthemispheric temporal and parietal cortex (i.e. brain lesions specific to deficits in memory, orientation and spatial abilities).

3.3.3.2 Procedure

For this within-subjects design each participant went through the same order of assessments for a total of approximately one hour. Patients were briefed about the experiment, signed a consent form and completed pen&paper and VE-based versions of a pointing task.

3.3.3.3 Task and materials

The pointing task included eight highly familiar target locations within the clinic which were not visible from the testing room (e.g. therapy rooms, clinic entrance, swimming pool). Targets were skipped if the location was unknown to the patient. For each target participants were given the target name and an A4-sheet of paper with a circle (diameter = 14cm) drawn in the center of the paper. The participants were instructed to imagine the middle of the circle as their current location. The circle was used to mark the direction towards the target. In addition, participants were asked about the linear distance towards the target and the difference in height between the floor at their current location and the floor at the target.

The computer-based pointing task was displayed on a 24-inch monitor, 60cm in front of the participant. The experiment was carried out with the same PC which was used to create the VEs (hexa-core CPU, 8GB memory, 2GB NVidia GTX 460 graphics card). The task environment was situated at a different location than the pen&paper test, but still familiar to all patients (Figure 20-D). Participants were not allowed to move away from their virtual location but were able to look around using the arrow keys of a keyboard. A red dot in the middle of the screen was explained to be the marker for pointing towards the target. The virtual viewpoint had to be moved by the participant so that the red dot pointed exactly towards the target. The participants were told to imagine the red dot to be similar to their index finger for pointing. After choosing a direction, the participants were asked again for distance and height difference for each target. Order of targets for both pointing tasks was identical for each participant. No feedback about pointing performance was given until the end of the test session.

3.3.3.4 Results

Angular deviation from target directions, distance and height judgments were averaged across all eight targets and subject to detailed analysis. Absolute values were used for angular deviation and height difference. A percentage measure was used for distance estimations (100% equals true distance). Initial Levene's tests revealed that the assumption of homogeneity of variances was violated for the angular deviation measure. Additionally, significant Shapiro-Wilk tests indicated that our data for all three variables were drawn from non-normally distributed populations. Consequently, non-parametric analyses were used throughout this study. Statistical equivalence can be established using the procedures of Tryon (2001) and Tryon and Lewis (2008). However, these methods rely on the calculation of t-values for confidence intervals. There is no established test procedure for non-parametric analysis of equivalence yet. Thus, Wilcoxon Signed-Rank tests were performed and effect sizes according to Morris and DeShon's equation 8 (2002) were calculated (Table 7). Non-significant differences and very small effect sizes for all three measures were hypothesized in order to support the equivalence of real and virtual mental maps.

Table 7. Nonparametric analyses and effect sizes for pointing task

Variables (unit)	Results					
	Mean ± STD, Median VE	Mean ± STD, Median Pen & Paper	Wilcoxon's z	p	N	Effect Size
Angular Deviation (degrees)	48.20 ± 24.27, 44.56	46.23 ± 36.56, 30.66	-0.76	0.442	27	0.07
Height Error (meters)	3.09 ± 1.92, 2.40	4.37 ± 5.16, 3.13	-1.48	0.137	26	0.43
Distance Estimation (%)	153.34 ± 215.22, 69.13	176.71 ± 198.70, 99.56	-1.20	0.228	26	0.21

3.3.4 Discussion

The employed analyses do not allow for firm conclusions about the equivalence of virtual and real environments, but non-significant differences and very small effect sizes are expected if both scenarios are very similar. All results show a very large variability. This can possibly be attributed to the inclusion of patients with a wide range of brain injuries. Future experiments need to selectively recruit patients with deficits in spatial orientation and memory deficits.

When patients pointed towards unseen targets, their judgment errors did not show significant differences in VEs and real environments. The resulting effect size is very small. Both groups committed similar errors and used similar strategies to find target directions. For example, when misjudging the first target, the following targets were shifted in accordance with the first pointing error (i.e. all targets misjudged by the same angle). When debriefing participants, the most common applied strategy was to mentally walk the path towards the target and update bearings during the mental wayfinding process. However, this strategy cannot be applied for distance judgments. Thus, most participants reported to have no strategy for distance or height judgments. Only few individuals used floor numbers and room heights as an indicator for total height difference between target and current location.

The small and medium effect sizes for distance and height estimations respectively, suggest that length is generally not judged equally in real and virtual scenarios. Moreover, several participants tremendously overestimated distances in both environments (i.e. >500%). In order to account for these outliers, medians were calculated for both groups (Table 7). Performance within the VE indicates that most participants underestimated the true distances towards all targets which is in line with previous studies (Furness & Henry, 1993; Witmer & Kline, 1998). Distance judgments in the real environment were mostly accurate. This difference cannot be explained by any obvious technical aspects of the VE. Patients were not able to walk through the real and VE in order to avoid the difference in visual and proprioceptive feedback during locomotion. Visual quality of the VEs was high when compared to earlier studies (Waller, 2000; Witmer & Kline, 1998). Unfortunately, no systematic assessment of judgment strategies was conducted. Hence, it cannot be determined whether participants used different strategies to judge virtual and real directions and distances.

Generally, mental maps of our sample of neurological patients appeared to be comparable between real environment and VE. However, dimensions of these maps differed considerably so that mental maps based on VEs were reported to be smaller compared to their real-world counterparts. Results of this study indicate that further studies are needed to explain the large variance in orientation performance. Such studies should aim to recruit patients with more strictly defined inclusion criteria, particularly addressing lesion size and location. Further, studies including healthy participants can be useful to establish a baseline of navigation performance in complex environments. Large environments were subject to experimental investigation in the past (Darken & Sibert, 1996; Waller, 2000; Witmer, et al., 1996). However, there is no clear evidence available that compares and explains navigation behavior and distance judgments in complex real and VEs. Experiments were either conducted in non-realistic environments (Darken & Sibert, 1996), did only address transfer of knowledge (Witmer, et al., 1996) or did not use comparable real-world environments and VEs for assessment of spatial knowledge (Waller, 2000).

As in previous studies (see chapters 3.1 and 3.2) the gathered data is of such complexity that further evaluations are necessary to fully benefit from the complex virtual scenarios which can be created with modern computer technology. Realizing the proposed workflow for VE-development builds the foundation for extending this approach to several applied fields. Yet, it is the quality and quantity of produced data which need to be further investigated before patient-centered VEs become accepted tools in rehabilitation.

3.3.5 Conclusion

Cognitive rehabilitation is dealing with the individual background and deficits of each patient. A context-sensitive approach to rehabilitation has already been proposed (Ylvisaker, 2003) and described in detail (see Chapter 1.1). VR technology has also gained momentum over the last decades (Rose, et al., 2005), especially in the area of rehabilitation. However, given the labor-intensity of creating VEs, the use of these tools is still not common in a context-sensitive rehabilitation approach. Creating individual VEs for each patient to provide a relevant training scenario requires the cost-effective, timely production of these tools. This current study tests the workflow to achieve realistic, to scale models of real environments (as described in Chapter 2.3). The development process has been demonstrated in a random sample of four complex virtual rooms, each taking less than six hours to complete. The finished environments were realistic enough for users to recognize and to associate with their real world counterparts. Virtual tasks in several cognitive domains have been implemented within the VEs for use in therapy or training. These tasks enable patients to train in meaningful environments that are based on the personal circumstances of the brain-injured patient. High fidelity and personal relevance of VEs are the basis for clinicians to make more accurate decisions about individual performance outside of the rehabilitation setting.

In summary, it has been demonstrated that user-centered VEs can be developed efficiently for use in context-sensitive rehabilitation. With a fast development process and sophisticated data collection capabilities, VEs provide a useful alternative to traditional tests and trainings. Future investigations need to explore the conceptualization of VE-based datasets and further evaluate the validity of collected data.

3.4 Experiment 4 - Single case trial

The final evaluation of this dissertation has been conducted between February 2011 and April 2011 at the Asklepios Rehabilitation Clinic, Schaufling, Germany. Patient HA was recruited for a single-case trial to evaluate the proposed workflow and usage of the embedded cognitive tasks. The patient was chosen for this trial based on several factors. Firstly, patient HA and her parents were willing to actively participate in this trial and provide access to the parents' home. Secondly, the patient's therapy goals were well-aligned with the content that can be provided by the VMT. By developing a model of the parents' house the patient could train to return home and complete tasks independently (e.g. preparing a meal). Lastly, the patient's health care plan was expected to pay for at least four weeks of neurologic rehabilitation which provided ample time for multiple training sessions.

It was the intention of this trial to apply the workflow and cognitive tasks during the day-to-day schedule of a brain-damaged patient (Aim 8). It was hypothesized that the VE relevant to the patient's therapy goals was developed within eight hours of development time (Hypothesis IV-a). Further, the cognitive training was expected to be adapted to HA's therapy goals (Aim 7). Lastly, the cognitive tasks were hypothesized to be of adequate difficulty for HA's training by not showing any floor or ceiling effect (Hypothesis II-a).

3.4.1 Subject

HA, 29 years of age, female, right-handed, shows extensive damage of both frontal-lobes and the left temporal lobe after a very severe traumatic brain injury in early 2010. HA's initial state is characterized by cerebral edema, brain stem contusion, and traumatic subdural hematoma. A bifrontal decompression craniotomy and hematoma relief are also reported alongside with initial mutism, severe autonomic crises, and spastic tetraparesis. After an initial coma of approximately four weeks HA was transferred to early rehabilitation about five weeks after the incident. Nine months after the accident HA was admitted to the neurology department of a rehabilitation clinic. Initially, HA was very passive and did not engage in conversations or social activities on her own. However, she adequately responded to questions and requests, albeit very slowly. Her motivation to practice and engage in therapies was low. HA spent most of the day in bed watching TV. Based on HA's tertiary education and professional experience, her premorbid intelligence was expected to be above average, especially for the cognitive domains of spatial reasoning and spatial memory. During initial neuropsychological tests, HA's verbal short-term and working memory were tested to be below the age norm. Further, HA's episodic memory several weeks before and after the traumatic brain injury was not accessible. Visual/spatial working memory was above average as tested by the WMS III's Block Span forward and backward (Wechsler, 1945/1997). HA was repeatedly able to reproduce sequences of seven items (blocks) in forward and reversed order. Spatial reasoning was assessed by the OPTT (Hegarty & Waller, 2004) in which she received a score in the first quartile of all 41 patients tested during the clinical trials in Germany (HA's average deviation 26.75 degrees, group median 35.46 degrees). In summary, patient HA’s abilities to remember events from the past and learn new events were compromised. Also, her motivation and social behavior were severely dysfunctional. Judging spatial relationships and layouts was one of her strengths and was still above average even with the traumatic brain injury.

3.4.2 Workflow

HA’s participation in this clinical trial was discussed with her, her family and the clinical team approximately two weeks after admission to the clinic. Her neurological rehabilitation was likely to be extended for at least eight weeks, so that prolonged training with the VMT was expected to be feasible. In order to promote independent living at home, HA’s training goal was to prepare a breakfast in her home environment. Consequently, location of relevant items and correct order of items were chosen as training tasks. After informed written consent was established, the home environment, particularly the kitchen, was measured and photographed during a home visit. Floor plans of the building were scanned and imported into Autodesk Design Review. The modeling process is depicted in Figure 21. The top left scene in Figure 21 shows the integration of the building’s floorplan into a SketchUp model. All subsequent scenes document the modeling of the building structure and interior details. The VMT was added to the VE as described in chapter 2.4.3. Twelve food and kitchen-related items were defined as target items. The actual development process took a total of eight hours. However, due to clinical obligations these eight hours of development were carried out over the course of a week. After this time period the environment was recognizable by the patient and the cognitive tasks were usable and ready for training sessions. However, after the initial development effort additional details and decorative items were added to the environment in order to showcase the finished application outside of the clinic environment, on websites and conferences. The application and a photograph of the real environment can be found in Figure 22.

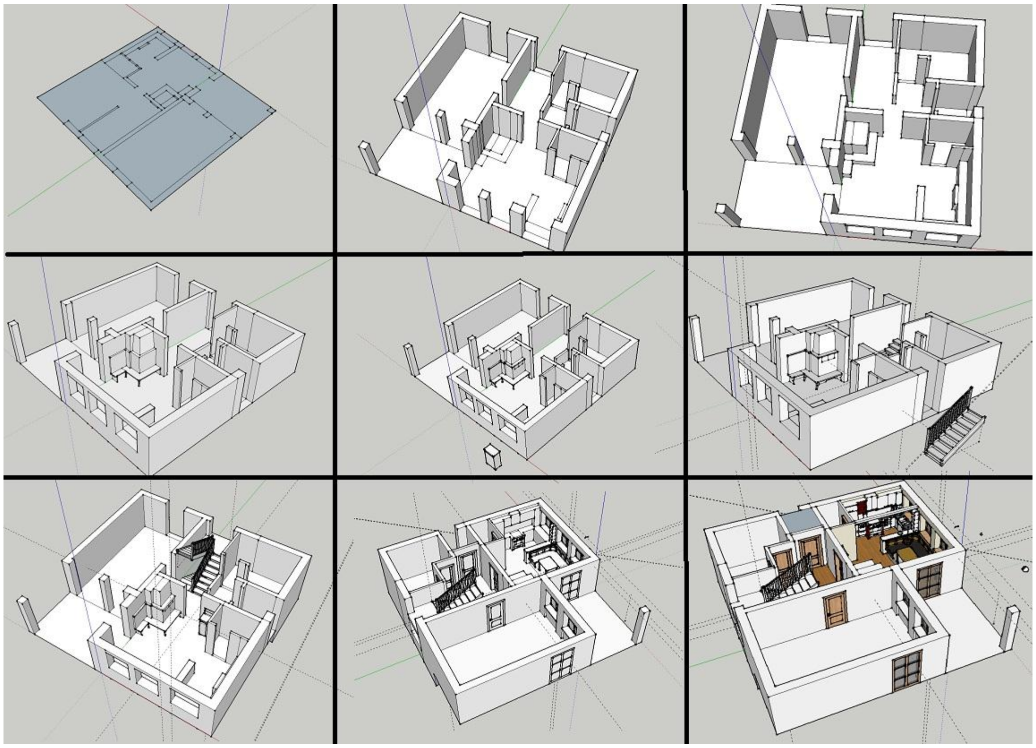

Figure 21. Development stages for 3D model within Google SketchUp

Figure 22. Real (left) and VE (right)

3.4.3 Training protocol

Prior to using the VMT, HA completed a series of neuropsychological tests identical to the protocol of Experiment two (see Chapter 3.2.2.2 for details). Due to the experimenter's commitment to clinical work outside of the single case study, the development process of the individual training environment was spread out over the course one week. While the virtual model of the home environment was still being developed, HA started using the standard version of the VMT. The VMT's standard version has been developed for evaluating the test with a larger clinical sample and is set in the virtual office room in which all assessments and training sessions of this trial took place. As soon as the patient's virtual home environment was fully functional both versions were used in parallel during training sessions. A total of nine training sessions were conducted over the course of four weeks. The first five sessions used the standard VMT and the remaining four sessions made use of the individualized VMT. Each session lasted between 30 and 60 minutes, consisting of 5 to 15 VMT trials per session (average 9.22). The frequency of training sessions varied considerably over the course of the study and depended on HA's therapy schedule. Once HA's mobility and motivation increased, physiotherapy sessions were of highest priority. Consequently, VMT training sessions were reduced to two sessions in the third week and one session during the last week of training. VMT sessions generally were scheduled towards the end of the daily therapy plan, because physiotherapy sessions occupied most of the morning and early afternoon. Due to time restraints the training was discontinued after the fourth week.

3.4.4 Results

As with HA's overall performance, VMT results varied substantially depending on her motivation, rehabilitative progress, and length of her daily therapy plan. The initial data presented here is taken from the trials using the VMT's standard version. Generally, a slight reduction in error scores for all trials involving viewpoint rotation and no rotation (Figure 23) and for trials involving 0 and 90 degree rotations (Figure 24) can be seen. However, given the low number of data points none of these differences justify the use of inferential statistics. No clear pattern emerged for trials with 180 degree viewpoint shifts. Further, no trend can be seen when trials involving four, five, six, and seven target items are analyzed (Figure 25). The average error score across all trials decreased from 0.296m for the first session to 0.215m during the ninth session. The error score represents the average offset of the placed target items from their original locations across all items of each trial, measured in meters.

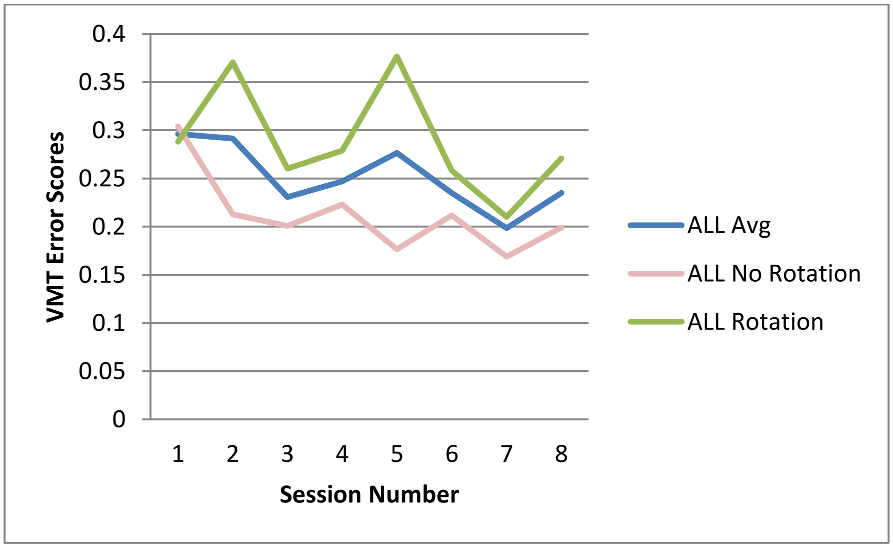

Figure 23. VMT error scores for rotation and non-rotation trials

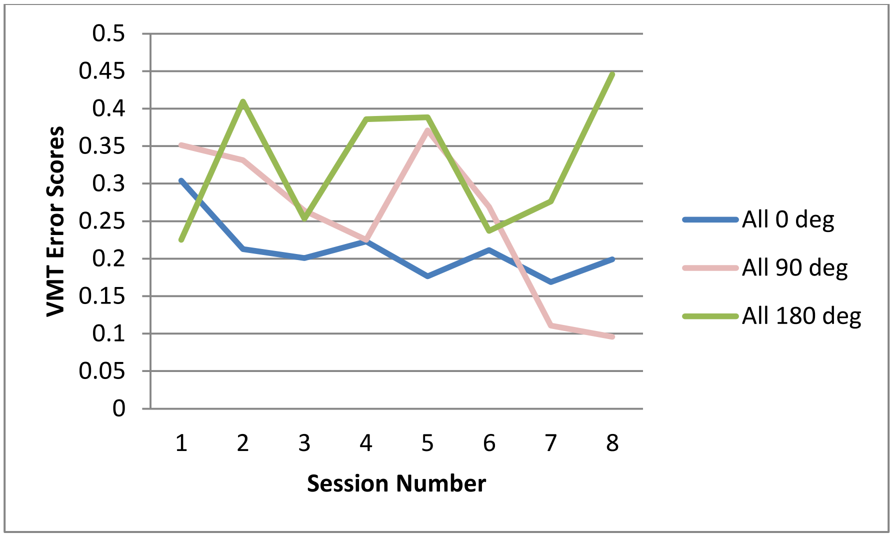

Figure 24. VMT error scores for viewpoint rotation conditions

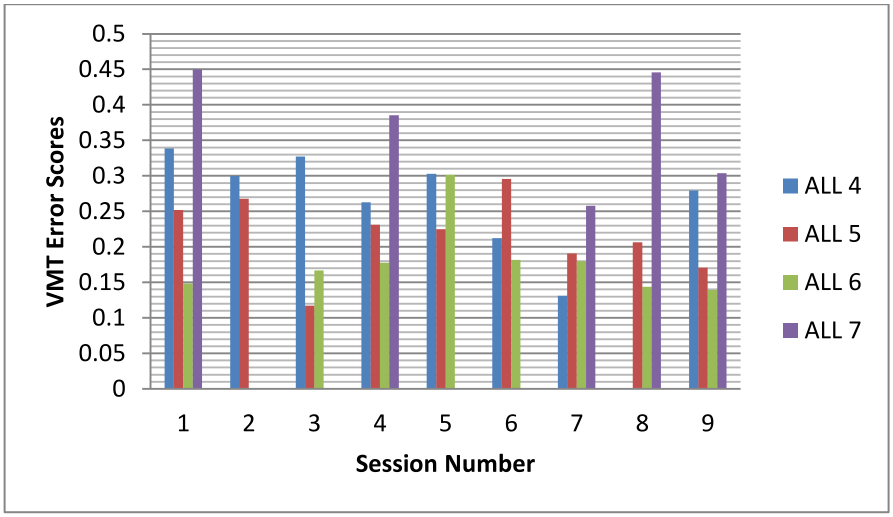

Figure 25. VMT error scores for different number of target items

The individualized VMT was only used during four training sessions so that statistical analyses are not justified. Session eight only consisted of one individualized VMT trial, as HA was too exhausted from earlier therapies to continue her training. The overall average error scores show a slight increase over the course of the four sessions. Average error scores are highest in session nine (0.268m; Figure 26). Comparing error scores for trials with different numbers of target items (Figure 27) and different perspective changes (Figure 28) does not yield conclusive results. It appears that performance in trials of low difficulty (i.e. 4 target items, no rotation) increases over time while trials of higher difficulty show large variability with the largest error scores during session nine.

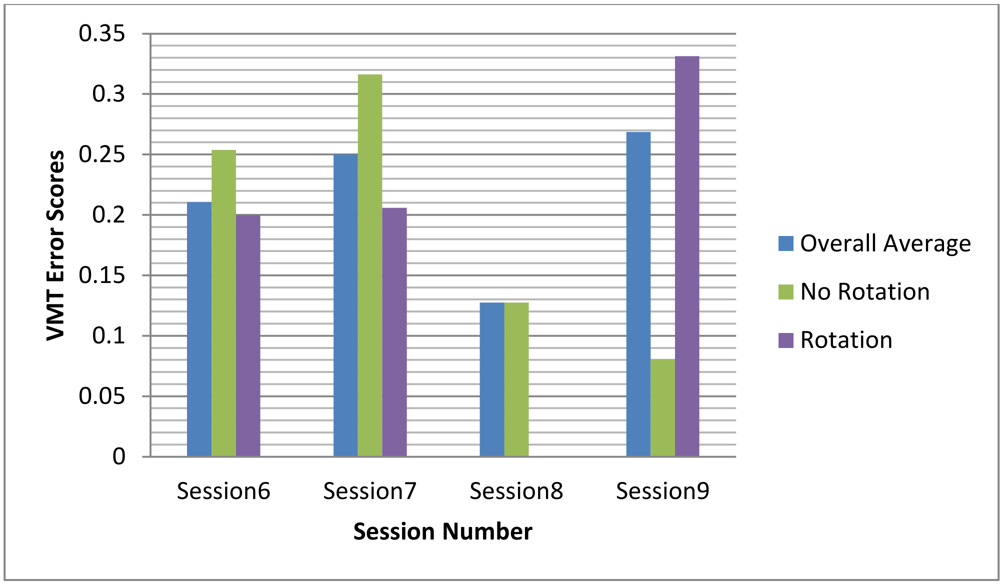

Figure 26. VMT (individualized) overall error scores

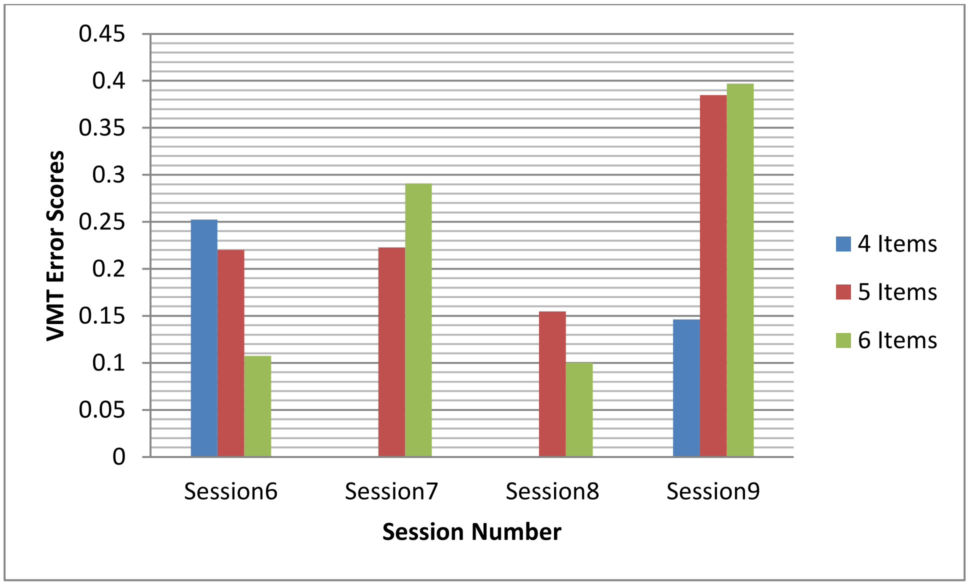

Figure 27. VMT (individualized) error scores for different number of target items

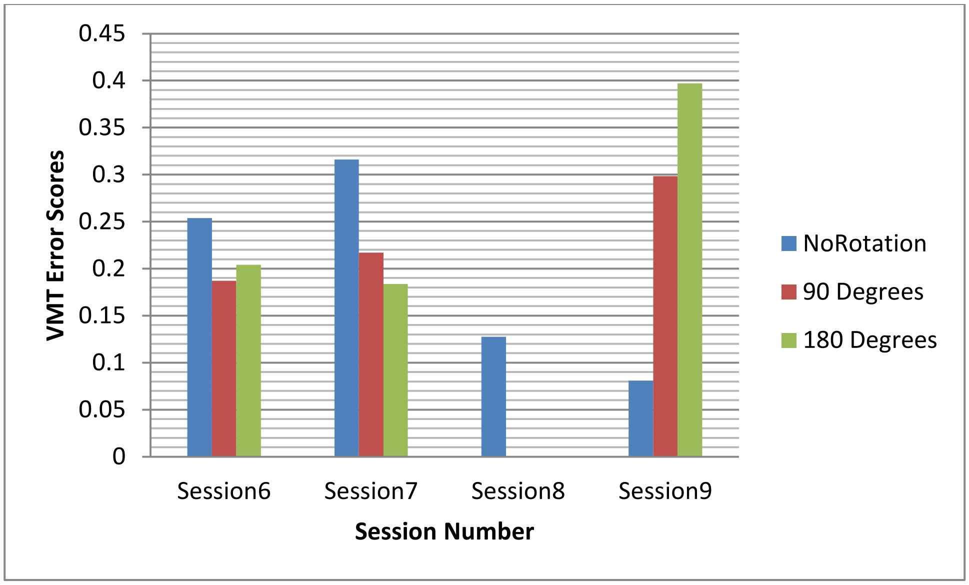

Figure 28. VMT (individualized) error scores for different rotation conditions

Results for an alternate use of the VMT indicate improvements over the course of three sessions. HA was asked to collect items from the kitchen table and put them in the order that is needed for meal preparation. Using six target items HA was able to correctly place the items for session seven and nine. During session eight five out of six items were placed correctly. More importantly, using seven target items HA achieved three, five and seven correctly placed items respectively (Figure 29). The order of items remained the same across sessions. While the constellation of all targets is relevant for the normal use of the VMT, targets only need to be put in order for this VMT variation. Because of the easier structure of this task, it is expected that these improvements are due to a learning effect.

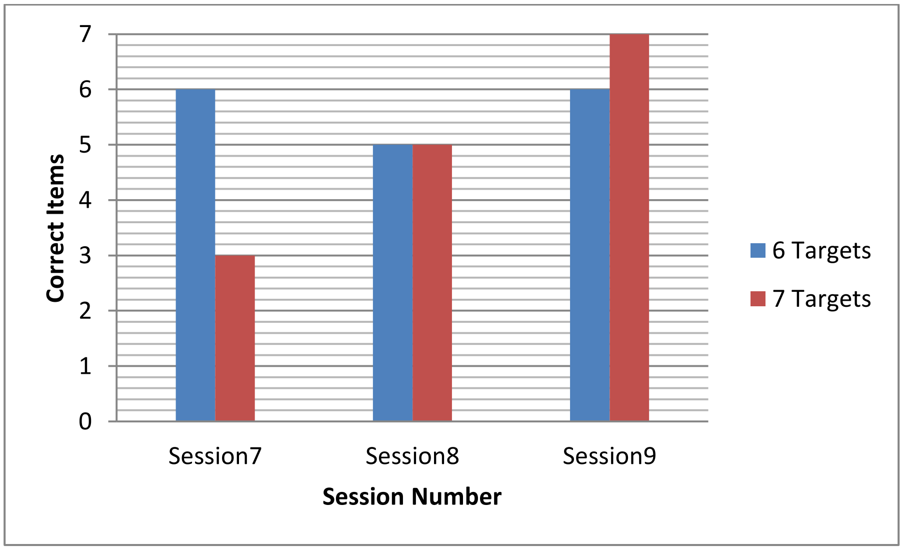

Figure 29. VMT (individualized) items correctly placed in order

Even though not enough data have been collected to statistically analyze HA's performance with the individualized VE, valuable insights about the paradigm have been gained during this single case trial. Most importantly, despite HA's passiveness, she was highly motivated to use the VMT. HA did not talk on her own nor showed any emotional reactions prior to starting her training. When first confronted with the individual version of the VMT, HA smiled and said that she knows the environment. She was able to identify furniture and several items in the VE. She also started to engage in conversations about the individualized application, its development and use.

Further, it was shown that the VMT is flexible enough to train the individual therapy goals of HA. Firstly, target items were scattered across the kitchen so that HA had to memorize their locations. Secondly, target items were aligned in a specific order and HA had to re-arrange the order in which the items would be used for specific tasks (e.g. preparing a meal). Thirdly, it is possible to scatter items around the environment and the user has to place them back to the same location repeatedly. This can be used to train patients to systematically organize items in their surroundings. Since HA already started to remember the location of items around the kitchen, the third training approach was not used.

3.4.5 Discussion

Despite the lack of data to statistically analyze HA's progress with the VMT over the course of the training, the development and use of the VMT can be considered successful. Despite the prolonged development process, the total time to develop the individual home environment (Figure 21 and Figure 22) was within the time limit set forth in Hypothesis IV-a. The virtual scenario was realistic enough to be recognized by HA. The relevance of the scene motivated her to use the training repeatedly and engage in conversations about the development and use of the application. Her training motivation and emotional responses when HA first saw the application is the most important outcome of this trial. However, her general increase in motivation and social activity throughout her rehabilitation might be best explained by the comprehensive therapy plan that HA received in many domains including physiotherapy and speech language therapy. The cognitive VMT training might have had a positive influence on the rehabilitation outcome, but currently this contribution cannot be quantified. Further trials with a more structured data collection are required to perform time-series analyses and evaluate the efficacy of the VE-based cognitive training.

Among the many components of the development cycle of such individualized cognitive training, the measurement and photographing of the actual environment were found to be the most labor-intensive. For future trials, a more automated approach of collecting information about the real environment is desirable. Specifically, the amount of time that the developer has to spend at the patient's home or workplace has to be reduced in order to minimize interference with the family's privacy or the employer's work routine. Preferably, an intuitive procedure would allow the patient's family or employer to capture the information themselves and send it to the developer for reconstruction of the environment's geometry and textures. Such procedure could involve the use of a depth-sensing camera (i.e. Microsoft Kinect) in order to capture the real scene appropriately.

4 DISCUSSION

The goal of this dissertation is the development and evaluation of cognitive tasks targeting navigation, orientation and spatial memory. The developed applications are based on elements of context-sensitive and process-specific rehabilitation by combining individualized virtual environments (VEs) and standardized cognitive tasks. The patient's individual context is incorporated through the use of meaningful, unique virtual environments VEs. Process-specificity is achieved by the use of cognitive tasks that target specific cognitive domains such as spatial and working memory. This development process is carried out by an experienced virtual reality (VR) developer who provides the individualized application to clinicians for daily use in rehabilitation. This training approach is expected to motivate patients, target relevant rehabilitation goals and provide valuable information to the clinician by accurately simulating situations in which the patient is expected to live independently.

The validity and feasibility of these tasks was evaluated during four experimental trials. During the initial stage of this dissertation experiment one tested the navigation ability of healthy adults in real and VEs. The remaining three trials were conducted at the Asklepios Rehabilitation Clinic, Schaufling, Germany. By recruiting patients with a wide range of neurological deficits it was intended to evaluate the developed tasks with users of varying cognitive and motor abilities. Due to the complexity and high variability of the collected data the quantitative outcomes of all trials did not produce conclusive evidence. However, the user feedback was very positive throughout and patients found the tasks easy to understand. Five patients were able to gain insight into their cognitive deficits by using the VMT during experiment two. This outcome is of central importance to this dissertation, because deficit awareness appears to be relevant for the successful outcome of cognitive rehabilitation (Ownsworth & Clare, 2006). If assessed patients do not realize the nature of their deficits and doubt the outcomes of the neuropsychological tests, their motivation for long-term cognitive training is expected to be low.

An additional relevant outcome was achieved while training a patient with severe traumatic brain injury. The individualized VE during experiment four encouraged patient HA to actively engage in conversations, show motivation towards continued use of the application and show positive emotions during practice sessions. Given the patient's passive state, lack of emotions and low motivation due to massive frontal lobe damage, these achievements are remarkable and suggest that the concept of individualized training tasks can be of high value for cognitive rehabilitation. A more detailed discussion of how the experimental trials address the aims and hypotheses of this dissertation can be found in the following chapter. Moreover, the thesis' limitations, potential, and future research will be discussed.

4.1 Aims

4.1.1 Aim 1

It is the primary aim of this thesis to develop a set of cognitive tasks targeting navigation ability, orientation ability, and spatial memory.

Each of the cognitive tasks has been implemented in the game engine Unity. They can be imported as unity-packages to be embedded into any VE that can be displayed in Unity. The VNT was the first development effort to be used with healthy adults. Its implementation in the virtual model of the Erskine building at the University of Canterbury was carried out over the course of three months. The task consists of a set of navigation targets which can be placed at any location within the VE. The user's ability to navigate from one target to the next is recorded and can be analyzed within the Unity editor.

The VMT is a cognitive task which combines elements of short-term memory and perspective-taking ability. The development of the task was planned as an iterative process which involved series of user testing of patients and therapists at the Asklepios Rehabilitation Clinic, Schaufling, Germany. The task can be customized within the Unity editor in order to create relevant training sessions for each patient.

The pointing task requires the user to point towards unseen targets. It was used as part of the VNT-trial and also in a separate study in which patients at Schaufling completed the task in the real clinic environment and its virtual counterpart. All three tasks can be used individually or in combination within a therapy session, depending on the patient's deficits. The tasks can be configured and set up in any VE within Unity prior to a session. Detailed descriptions of each task can be found in chapter 2.4. By completing the development of the tasks and evolving each application through continuous testing with healthy participants and neurological patients, this aim has been successfully achieved.

4.1.2 Aim 2

It is an additional aim to assess each cognitive task's validity in an experimental trial.

Four experiments have been carried out as part of this thesis. Experiment one was intended to assess navigation performance of healthy adults in a complex real-world environment and a virtual to-scale copy of the scene. This trial involved the development and testing of the VNT and a first prototype of the pointing task. Healthy adults were recruited for this initial data collection due to the prototypical nature of the VNT assessment, the exhausting study protocol and the limited availability of neurological patients in Christchurch, New Zealand. During the initial stage of this dissertation, the individualization of each participant's environment was not a major goal yet, so that each participant was exposed to the same virtual/real scenario. By showing that navigation performance is equivalent in real and virtual spaces, this experiment was expected to provide evidence that the VNT can be utilized in a clinical setting. The results of this trial indicate that navigation through complex environments cannot easily be assessed with traditional measures of spatial abilities. Further trials need to evaluate the influence of environmental difficulty parameters (i.e. visibility of landmarks, familiarity of the environment, number of alternate routes, etc.) and navigation choices on the user's performance. Also, an evaluation of navigation performance with brain-injured patients is necessary to show the VNT's relevance for clinical decision-making.

Experiments two, three and four were conducted at the Asklepios Clinic, Schaufling, Germany. The purpose of the second and third experiment was to test the VMT and pointing task with a diverse sample of neurological patients. The intention was to receive patient feedback about task difficulty and user interaction while also collecting data about the validity of the cognitive tasks. While the outcome of the validity evaluations did not provide clear results throughout, the concept of the VMT did show significant correlations with traditional measures of short-term memory. Further evaluations with more homogeneous samples of neurological patients (e.g. lesion to right temporal/parietal cortex) are expected to provide additional evidence for the test's validity. More importantly, experiment two clearly showed the importance of the VMT's transparent nature that closely resembles tasks that patients recognize from their daily life. A total of five participants were able to gain insight into their cognitive deficits by using the VMT.

Experiment three was successful in replicating results of previous experiments in which a pointing task resulted in participants underestimating virtual distances as compared to actual real-world measurements. This experiment also demonstrated that VEs created with the proposed workflow (chapter 2.3) are realistic and resemble real-world environments accurately. Healthy participants recognized the environments in at least 88% of all cases.

Experiment four applied the previously tested cognitive assessments to a brain-injured patient. This single-case trial was planned to demonstrate that an individualized VE can be created within the proposed timeframe of eight hours and that task content and difficulty are appropriate for clinical use with a patient. While the development of the VE was interrupted by clinical obligations of the experimenter, the total development time was within the set time limit. Patient HA's therapy goal to return home and live independently without the need for assistance in activities of daily life was integrated into the VMT. VNT and pointing task were not relevant for the patient, as navigation and spatial orientation were left unimpaired after HA's severe traumatic brain injury. Because of the unpredictable nature of neurological rehabilitation, patient HA's focus shifted towards physio- and occupational therapy as time progressed. Therefore, only few VMT sessions were conducted and not enough data was gathered to analyze the results statistically. Nonetheless, HA's motivation and positive feedback are a good basis for future trials of this individualized rehabilitation concept. Taken together, each task was tested in experimental trials with healthy adults and neurological patients. While the outcome of the validity assessments remains inconclusive for most outcomes, the qualitative feedback from participants is of high importance and lends support for further expansion and evaluation of individualized rehabilitation in VEs. The second aim of this dissertation has been fulfilled.

4.1.3 Aim 3

It is aimed to develop an optimized workflow for creating individualized VEs.

During the course of this dissertation an effective workflow for the development of VEs has been established. Several game engines and applications for 3D modeling were tested for their compatibility and their ability to rapidly create realistic virtual scenarios. The game engine Unity and Google SketchUp were chosen for their ease of use and active user community. It has to be noted that this workflow is based on the personal preferences and skills of the developer. These choices may change over time with the release of software updates or new applications. Each software and development choice as described in chapter 2 can be adjusted to match the developer's skillset and preferences. Further, it is acknowledged that these procedures do not constitute innovative concepts and are for the most part common knowledge for game and software developers. However, the proposed workflow is intended to serve as a suggestion for researchers and developers of clinical applications to minimize development time and costs. After the workflow was presented at the International Conference on Virtual Rehabilitation 2010, Zurich, Switzerland (Koenig, Dünser, et al., 2011), several research groups showed interest in this development concept and have adopted the described procedures into their projects (Sangani et al., 2012). Hence, aim 3 has been accomplished.

4.1.4 Aim 4

It is an aim of this thesis to integrate each cognitive task in a meaningful VE.

The game engine Unity provides the option to import Unity packages into a project. Each cognitive task has been developed within Unity and exported into such package. Whenever an individual VE is created for training a patient, the appropriate task only needs to be imported into this scene and set up as described in chapter 2.4. This modular approach ensures that only minimal time is spent on the integration of VEs and cognitive tasks. The integration was demonstrated for each experimental trial of this dissertation. Hence, this aim has been achieved.

4.1.5 Aim 5

It is an aim to test the efficiency of the development process of the VEs:

a. *in a controlled setting.*

In experiment three the development process was evaluated by randomly choosing four rooms from the Asklepios Rehabilitation Clinic, Schaufling, Germany. Each room was modeled to-scale in less than six hours by an experienced developer. Development times can vary depending on the developer's experience with tools and workflows. For the purpose of this thesis the VEs were detailed and realistic representations of the actual real-world scenarios. The sufficient quality of these models was demonstrated by a survey in which clinic staff recognized the environments in at least 88% of all cases.

b. in a clinical, patient-centered setting.

Experiment four involved the measurement, modeling and clinical usage of an individualized VE. Patient HA's kitchen was modeled to scale and used for the training of spatial memory. During a home visit the kitchen was measured and photos and videos of the environment were taken. The modeling process took a total of about eight hours when using the described workflow with an experienced developer. However, the development was spread out over the course of one week due to clinical obligations of the developer. Additional time was invested once the application was fully functional to add further details to the environment in order to showcase the scene outside of the clinic. Considering the initial 8-hour development effort, this aim is considered to be completed. In the future, the workflow and development tools would ideally be adjusted to empower the clinician to set up characteristics of the VE and make changes to tasks and collected data. As of now an experienced developer is required to fulfill this role.

4.1.6 Aim 6

It is aimed to apply the embedded cognitive tasks throughout the neurological rehabilitation of a brain-injured patient.

Patient HA suffered from a severe traumatic brain injury with substantial damage to bilateral frontal lobes and left temporal lobe. The patient's rehabilitation goal was to return home and live independently without the need to receive help for activities of daily living. Cooking a meal was identified as an important task that requires planning, spatial memory and working memory. After the kitchen environment was modeled in sufficient detail, kitchen utensils and groceries were placed in the actual locations in which they also could be found in the real kitchen. These items were defined as target objects and integrated into the VMT. HA used the VMT as it was originally intended by moving targets to previously learned locations either with or without perspective changes. Also, HA trained the order in which targets were needed for meal preparation by picking the items up and placing them next to each other in the correct order. More detailed information about the use of the VMT can be found in chapter 3.4.4. VNT and pointing task were not used during this single-case trial as patient HA's orientation and navigation ability were unimpaired. Due to the fact that the focus of HA's rehabilitation shifted towards physical and occupational therapy, the number of training sessions using the individualized VMT is low. Further prolonged use of the VMT with brain-injured patients is necessary to draw conclusions about the test's feasibility for clinical use. In addition, a randomized controlled trial is needed to establish the VMT's potential for improving cognitive abilities during neurological rehabilitation. Aim six has only been partially accomplished.

4.1.7 Aim 7

It is an aim to use the VMT to accommodate a patient's individual therapy goal.

The VMT has been used in a single case trial (Chapter 3.4) during which the application has been adapted to the patient's therapy goals. The main goal for patient HA was to regain independence in functional tasks such as personal hygiene, dressing, cooking and eating. While most goals were concerned with the regaining of motor function, meal preparation was selected as a goal for cognitive rehabilitation. Making a meal involves remembering where relevant items in the kitchen are stored. They also have to be used in the correct order (e.g. bread is needed before butter can be used). HA used the VMT to train to remember the location and use of kitchen-related items in the environment. The individualized task was set up to replicate the actual location of the target items which HA had to remember. A total of three different ways to utilize the VMT to reach HA's therapy goals were found. Only two of the three tasks were actually used in therapy, as HA started to remember how the kitchen was organized on her own. Results of this trial show first support to indicate that the VMT is flexible enough to adapt to a patient's individual therapy goals. Similar flexibility can be expected from the navigation task as navigation targets can be adapted to the patient's needs. However, since patient HA's training did not include navigation or spatial orientation, no such conclusion can be drawn without further clinical trials. Given the heterogeneity of neurological deficits that are to be expected in cognitive rehabilitation these results have to be replicated across a larger sample of patients before conclusions about widespread use of this individualized rehabilitation approach are possible.

4.1.8 Aim 8

It is an aim to integrate the proposed workflow into the rehabilitation routine of a brain-injured patient.

Patient HA suffered from a severe traumatic brain injury which caused extensive damage to both frontal lobes and the left-sided temporal lobe. The initial therapy goal for patient HA was to live independently at home without requiring assistance in basic activities of daily living (e.g. eating, personal hygiene). During experiment four the VMT was used to train HA's ability to prepare a meal. Navigation and orientation ability were not deficient so that VNT and pointing task were not relevant for the patient's rehabilitation. A home visit was scheduled after the study protocol was discussed with the patient and parents and informed consent was established. During this two-hour visit the real environment was measured and photographed. Based on this information the virtual scenario was created over the course of one week. The actual development time was approximately eight hours, but due to clinical obligations of the experimenter the development process was delayed. The detailed development process is described in chapter 3.4.2. Due to the delayed development this aim was only partially accomplished. However, in actual clinical use this limitation would be irrelevant as the developer creating the VE would not be involved in clinical work. Future trials need to replicate this workflow integration with a wider range of brain-injured patients.

4.2 Hypotheses

4.2.1 Hypothesis I

Cognitive tasks integrated into the VEs are expected to target specific cognitive processes (process-specificity).

a. *The VNT is hypothesized to show equivalent outcomes of navigation measures as compared to a real-world navigation task.*

The results of both evaluation studies (Chapter 3.1 and 3.2) do not allow for strong conclusions about the validity of the proposed cognitive tasks. The navigation task's results show large variability for a sample of 31 healthy participants navigating through a real-world complex building and its virtual counterpart. Several of the study's assessed variables are affected by difficulties of participants interacting with the VE. An analysis based on Tryon's approach of equivalence testing (2001) between participants navigating through real and VEs does not reveal conclusive evidence for both scenarios being equivalent. However, the absence of large effect sizes and statistically significant differences between both groups suggest that the navigation experience in real and VE are somewhat similar. Due to the absence of conclusive, significant results, this hypothesis cannot be supported. However, these initial results suggest that additional studies assessing navigation performance in complex VEs are justified. No prior studies were found that evaluate equivalence of navigation behavior in healthy and brain-injured participants. It can be expected that cognitive performance, and thus also navigation performance, is markedly impaired in brain-injured patients (Rao, Jackson, & Howard, 1999; Schretlen & Shapiro, 2003). However, it is currently unknown how lesion locations (e.g. frontal lobe damage affecting navigation strategies) and severity of injuries might differently affect navigation performance in real and virtual environments and whether equivalence of performance in real and virtual environments applies to this population. However, many studies have been conducted about wayfinding performance in brain-injured patients (e.g. Livingstone & Skelton, 2007; van Asselen et al., 2006). A future study needs to reveal how memory deficits, planning deficits or impairments in spatial processing and imagery affect performance in VEs.

b. *The VNT is predicted to significantly correlate with pencil and paper measures of spatial abilities.*

All VNT outcome measures were not significantly correlated with pencil and paper tests of spatial abilities. Only errors in an additional task, the floor plan task, were significantly correlated with the Card Rotation Test (CRT). Poor performance in the CRT was associated with larger number of errors for drawing in targets and current location on a floor plan of the Erskine building. The lack of convergent validity with measures of spatial abilities suggests that navigation performance in complex environments cannot be explained by traditional assessment methods. It is expected that such difficult scenarios with many possible routes towards the target location require the integration of multiple cognitive abilities such as working memory, planning, focused attention, and problem solving. Since the assessed environment was a public building (campus building) with many people and distracting stimuli, the difficulty of the task was much higher than traditional navigation tasks (Koh, et al., 1999; Witmer, et al., 1996). Further evaluations will be required to reveal the relationship between navigation performance in complex environments, difficulty parameters of the task/environment, and cognitive abilities. This hypothesis has not been supported.

c. *The VMT is hypothesized to significantly correlate with established neuropsychological tests that assess spatial memory.*

The VMT's evaluation shows promising results for its future use in a clinical context. Study outcomes suggest that the test's scores are significantly correlated with results of the Rey-Osterrieth Complex Figure Test which is a test for spatial/visual memory. Based on these results, the hypothesis is supported.

d. *The VMT is hypothesized to significantly correlate with established neuropsychological tests that assess spatial abilities.*

No correlations have been found with neuropsychological tests of spatial abilities (e.g. perspective taking and mental rotations) and working memory. Though, the conceptualization of test results plays an important role in these interpretations. Results of the VMT are much more complex than the correct-incorrect dichotomy of traditional test items. Furthermore, the study's clinical sample was very heterogeneous in regards to brain injuries and cognitive abilities. Consequently, the validity of the VMT has to be re-assessed with homogeneous samples of relevant target populations as well as healthy participants of several age ranges.

e. *It is predicted that the VMT does not show significant correlations with cognitive tests of domains unrelated to the VMT.*

The VMT did not correlate significantly with the D2 test of attention. This result is an indicator for the VMT's divergent validity. However, since the test also did not correlate with measures of spatial abilities, the absence of correlations with standard pencil and paper tests might only be due to the complex nature of the VMT. Further evaluation is required to confirm whether the results of the proposed tasks are unrelated to the outcomes of traditional neuropsychological assessments. Moreover, future studies need to assess whether the proposed tests provide information that is more relevant for the everyday performance of brain-injured individuals.

f. *The pointing task is hypothesized to show equivalent results in a real environment and its virtual counterpart.*

Despite the similarity of navigation parameters in both groups of experiment one, a consistent difference in distance judgment emerged across all studies. During the navigation trial and the mental maps task (Chapter 3.3.3) participants consistently underestimated distances within VEs. These results are in agreement with existing studies (Witmer & Kline, 1998) and have to be taken into account when distances are of interest for cognitive assessment and training.

The conclusions from the previous hypotheses suggest that the validity of the modular cognitive tasks requires further evaluation. Throughout this dissertation many insights regarding task conceptualization, study design and data capture have been gained. Consequently, the presented preliminary results are an excellent starting point to further develop the set of cognitive tasks in future studies. Conclusions based on current results need to be drawn carefully while taking into account the applications' prototypical nature, diverse sample and the resulting data's variability. Based on the results of the completed evaluation studies, Hypothesis-I cannot be accepted. Most comparisons to traditional neuropsychological assessments do not support the validity of the proposed tasks. Though, as previously indicated, there are large conceptual differences between the proposed complex tasks and pencil and paper tests.

4.2.2 Hypothesis II

The proposed applications are predicted to be flexible enough to meet the changing demands of a patient's neurological rehabilitation (context-sensitivity).

a. *The VMT is expected to be used throughout a patient's neurological rehabilitation without the occurrence of a floor or ceiling effect.*

During a single case study (Chapter 3.4) the VMT has been used to train patient HA for a total of nine sessions over the course of four weeks. For the purpose of this study 30 different arrangements of target items were created. Nine of these setups used the individualized kitchen environment of patient HA, all others were based on the standard version of the VMT. The arrangements covered tasks with four to seven target items and 0, 90, and 180 degree viewpoint shifts. This setup provided a wide range of task difficulties and was expected to be sufficient for training sessions throughout a patient's neurologic rehabilitation. Over the course of patient HA's training sessions no clear performance pattern emerged. Task performance varied considerably across trials but HA always appeared to be challenged appropriately. In easier trials HA smiled while reporting a sense of accomplishment. Harder trials (i.e. six to seven targets, 180 degree perspective change) left the patient motivated to try again without being discouraging. After the second week of cognitive training patient HA's main focus shifted towards the regaining of motor abilities, so that only a low frequency of VMT trials was accomplished.

Considering the absence of any floor or ceiling effects, task difficulty has been found to be adequate throughout the four-week training period. However, prolonged task exposure in future clinical trials in which patients partake in high-frequency training will be necessary to substantiate evidence for accepting this hypothesis in a wider population of cognitively impaired individuals. Within the scope set forth in this dissertation, Hypothesis II-a has been supported. The VMT is flexible enough to relate to the dynamic nature of a patient's neurologic rehabilitation without the occurrence of floor or ceiling effects.

4.2.3 Hypothesis III

The workflow for creating the proposed individualized training is expected to be suitable to create realistic, high-fidelity environments with a high degree of ecological relevance.

a. *It is hypothesized that developed VEs show high recognition rates by users.*

Experiment three evaluated the quality and realism of four VEs that were modeled after four randomly selected rooms at the Asklepios Rehabilitation Clinic, Schaufling, Germany. Staff members were asked to identify screenshots of the VEs and also to recognize the locations of the VEs out of a set of 20 photographs of rooms throughout the clinic. While identification rates ranged between 56% and 98%, the recognition rates were above 88% for all four rooms. These results suggest that most participants are able to recognize the VEs and associate them with the respective real-world environments. Consequently, VEs created by the proposed workflow are of sufficient fidelity and visual quality to represent real-world locations. Hypothesis III-a has been confirmed. Future studies can potentially address the effect of visual quality and visual details on the outcome of cognitive tasks and recognition rates. However, the difference in development efforts between highly detailed and low-detailed environments is expected to be minimal. Hence, a highly detailed environment of high visual quality appears to be the preferred choice.

b. *Cognitive tasks are expected to be transparent and easy to understand by users.*

During experiment two the VMT was shown to be effective for patients who were denying their cognitive deficits. A total of five patients were able to gain an understanding of their deficits due to the transparent, easy-to-understand nature of the VMT. When debriefing the patients, they reported that they were able to relate to the task and understand its relevance for their daily life. Feedback from users in all other experiments was mostly positive. Only during the first experiment three participants had to withdraw from the study due to symptoms of simulator sickness. The VNT was the first prototypical development effort during this dissertation. Issues with display distortion on large projection screens and the mouse-based user interaction caused problems for some participants. However, no evaluation surveys were used to assess the usability of the developed applications. All feedback was gathered verbally after each training or testing session. Based on the quantitative feedback from the conducted experiments, Hypothesis III-b can be partially accepted. An additional trial will be necessary to further explore the ecological validity and usability of the proposed application and to substantiate this hypothesis.

4.2.4 Hypothesis IV

The workflow for creating the proposed individualized training is expected to be effective enough for integration into the daily routine of a rehabilitation clinic.

a. *Each functional training environment should be created in less than one working day (i.e. eight hours of development).*

The evaluation of the proposed workflow included the modeling of four random rooms of a large-scale hospital building (Chapter 3.3.1). Each room was measured, photographed and modeled with Google SketchUp by an expert developer within less than six hours. During an actual clinical trial the development process also involves travel to the patient's home environment or workplace. As a consequence, a functional training environment can often be ready for use within one or two working days, depending on the availability and communication with the patient's family or work superiors. Also, high familiarity and experience of the developer with an effective workflow and specific tools is required. A wide range of tools and procedures is available to achieve the desired outcome. However, the exact development time of a functional VE might vary depending on the developer's skill and used tools. Within the scope of this thesis the development process was shown to be adequate for use in the context of a rehabilitation clinic. As long as a single room is sufficient for a patient's cognitive training and expertise with the proposed workflow or a similar workflow can be assumed, Hypothesis IV-a can be accepted. If more rooms or highly-detailed environments are of importance for a patient to regain independence in daily activities, the workflow for creating VEs has to be further optimized.

4.3 Limitations and future work

The prototypical state of the proposed application places several restrictions on the interpretation of the collected data and the usage of the rehabilitation system. However, each of these restrictions opens up opportunities for future development and expansion of the current system. The following chapter discusses limitations of the proposed rehabilitation framework and how they can be addressed in the future. Further, possible extensions of the current system are portrayed.

4.3.1 Workflow

As previously demonstrated the proposed interactive VEs can be created within six hours as long as the real scenario is limited to a single room (Chapter 3.3.1). However, the current workflow is still labor-intensive and poses several restrictions on the everyday use of the proposed system in a clinical context. A single virtual room is not always going to be sufficient for cognitive training, especially when navigation and spatial orientation are of interest. Therefore, the development time will be a multiple of the estimated six hours, depending on the number and size of relevant rooms. A further bottleneck for system development is the acquisition of measurements, photographs, and video footage of the real environment. The developer has to arrange for a meeting with the patient's family or employer, travel to the real environment, and spend at least two to three hours collecting information about the relevant scenario. This workflow integrates families or employers directly into the rehabilitation process which can be considered an advantage over traditional therapy protocols. However, the information collection essentially poses an intrusion into the family's privacy or the employer's daily work routine. Consequently, alternatives to the current workflow should be evaluated that automate this process or minimize the time spent at the relevant real environment. One option is to actively involve families and employers in this information acquisition. Unfortunately, without any knowledge of the 3D modeling process it is challenging to take adequate photographs and videos of the environment. Hence, the family or employer would need substantial guidance to collect measurements and photos themselves. From the 3D modeling experience gained during this dissertation it turned out to be much more difficult to model a virtual space without ever having seen the real space. While a video walkthrough of the actual environment is helpful for creating the virtual model, the developer's experience of personally walking through the real environment can help tremendously when creating realistic virtual scenes.

The use of laser scanners[34] or depth-cameras (Izadi et al., 2011) needs to be considered as an additional option for creating 3D models of real environments. However, laser scanners are often expensive and commonly produce data for computer-aided design (CAD) applications. Geometry in CAD-formats (e.g. .DWG or

---

[34] Faro GmbH – http://www.faro.com

.DXF formats) is very complex and not optimized for real-time 3D applications. Hence, the CAD-models would need to be exported to a 3D modeling application and simplified significantly to make them a viable option for real-time rendering. This complex procedure would offset the initial time savings and result in a much more costly workflow.

A cost-effective alternative for real-time scanning of complex geometry has been demonstrated using a Microsoft Kinect (Izadi, et al., 2011). This concept not only provides an option for the rapid production of realistic 3D models and VEs, it could essentially change the way we interact with real and virtual spaces. If the Microsoft Kinect can capture the environment dynamically in real-time, cognitive tasks could eventually be integrated directly with the information that the depth-sensing camera provides on-the-fly. The user would then interact with the real item via see-through optics instead of a virtual representation of items and environments on a computer screen. Alternatively, the captured 3D geometry could be exported and used in a traditional setup at the hospital as it has been described in previous chapters. Given the rapid development of creative projects using the Microsoft Kinect[35], these scenarios will probably be feasible in the not-so-distant future.

4.3.2 User interaction

User interaction was considered an integral part of system development. Prior to using the VMT, navigation task, and pointing task in experimental trials, several user interfaces were tested. Using a simple three-button-mouse emerged as the best option for virtual navigation, even though some users still struggled to effectively move through the virtual scenarios. For the VMT and pointing tasks a shared interface between therapist and patient was employed. Keyboard and mouse were redundantly mapped with all commands required to use the applications. The patient was able to make selections and move items with the mouse while the therapist had the option to support the patient via keyboard whenever necessary. Some patients struggled to understand the concept of pointing towards unseen targets using the computer mouse and a red target pointer on the computer screen. In such cases additional instructions were necessary. However, almost all participants were able to use the mouse-based interface with only little assistance. However, patients with hemiplegia, tremor or other motor disabilities were very inaccurate with mouse movements, especially when attempting to drag and drop small target items. Further evaluations of alternate interfaces will have to be conducted in the future to avoid introducing this source of error into the outcome of the cognitive tasks. For example, the use of a modified controller for discrete button presses (e.g. game pad), body-/hand-tracking through Microsoft Kinect or web-cameras, or brain-computer interfaces such as the Emotiv Epoc[36] might be viable alternatives to the currently used interface. Ease of use for older patients without computer experience and patients with motor disabilities need to be considered. Most importantly, the users' experiences need to be quantified in order to make

---

[35] Kinecthacks – www.kinecthacks.net

[36] Emotiv Epoc - http://www.emotiv.com/

well-informed decisions about which interfaces to use in future iterations of the proposed rehabilitation framework.

Displaying the VEs was a major concern for the evaluation of the navigation task (Chapter 3.1). The experimental protocol involved a three-screen back projection system with a field-of-view of 120 degrees. Distortion of the user's viewpoint through the virtual camera and problems with using the computer mouse for navigation caused three users to report symptoms of simulator sickness and several other participants to show mild symptoms of simulator sickness. However, the unique display setup of this experimental protocol was of no relevance for the everyday use of the proposed rehabilitation system. The cognitive tasks are intended for use on a high-performance Windows-based PC with a single monitor. None of the neurological patients tested in Experiments 2-4 (Chapters 3.2, 3.3, 3.4) reported any major signs of simulator sickness when using such desktop setup. However, when developing applications for brain-injured individuals, additional care regarding the exposure to interactive VEs and provocative visual stimuli is required. Even though the desktop setup is of low immersiveness compared to head-mounted displays or large projection screens, patients with traumatic brain injuries or epilepsy require special consideration during development and usage of such system. Stanney's (2002, pp. 721-730) recommended protocol still represents the most recent standard for exposing participants to VEs and should be used when designing future experiments. It contains a list of guidelines for the development and exposure of users to a VE system. A more intuitive and restricted input scheme can also contribute to the safety and comfort of the users (Stanney, et al., 2002).

In summary, the proposed framework has been tested with a wide variety of healthy and brain-injured patients. The feedback gained during these trials has been invaluable and improved each of the cognitive tests during their respective experimental trials. During each study a wide range of users with different levels of abilities were chosen. This was a limiting factor to the validity evaluation of the developed tests. On the contrary, the verbal feedback from patients suggests that the proposed system is applicable and usable with a heterogeneous patient population. While no systematic user feedback was collected, patients commented on user interface, task mechanics, motivation and task transparency. This anecdotal feedback suggests that the developed tools are suitable for future validity evaluations and randomized controlled clinical trials.

4.3.3 Evaluation design and analysis

Two conceptual issues were of critical importance across all conducted evaluation studies. Firstly, a primary goal of all evaluations was to show that cognitive processes are similar when the user is exposed to VEs as opposed to experiencing the real world. This by itself poses a problem, because most statistical analyses done in the past decades are concerned about null hypothesis significance testing in which the experimenters are looking to find significant differences between groups or testing conditions. However, the absence of a significant difference does not provide evidence for the null hypothesis (i.e. no difference between the means of the tested samples/conditions) to be true (Nickerson, 2000). Tryon (2001) and Tryon and Lewis (2008) propose a procedure to statistically test for equivalence between two groups. However, their analysis is based on the overlapping range of inferential (shortened) confidence intervals. This range is compared to an interval (delta) that is considered to be of no consequence regarding the difference of both groups. Unfortunately, no guidelines exist that indicate what information delta needs to be based on. According to the author (personal communication with W. Tryon, January 29, 2010) the range of delta needs to be determined on substantive grounds and should be based on its practical consequences, a reasoned argument, and clinical experience. This however leaves much room for subjectivity and results in case-by-case decisions. In Experiment 1 (Chapter 3.1), for which an equivalence analysis has been conducted, one wrong turn or incorrect decisions can already result in a patient getting lost and not finding a target. Should therefore delta be set to an error score of zero? What is the relevance of a wrong navigation decision? These questions essentially lead to the second fundamental issue concerning the conducted experiments.

The data output of the developed applications exceeds the complexity of most traditional neuropsychological tests. The navigation task records the exact path of the user so that each individual decision and the complete route are stored and available for analysis. With such data complexity questions arise about how this information can be used for clinical decision-making. Does it matter whether the patient arrives at the target on a suboptimal route or after walking in circles? It is critical to ask how such differentiated information can add to the current practice of using established questionnaires and neuropsychological tests. Eventually, the information provided by the proposed cognitive tasks needs to be used by clinicians for whom the output of a lengthy log file of spatial coordinates and time stamps is of little use. Recording of navigation paths for visual analysis has already been established as a valuable tool (Werner, Rabinowitz, Klinger, Korczyn, & Josman, 2009). However, there are still no established standards of how such information can be useful in a clinical context. Eventually, a compromise has to be found between simplifying the available information for everyday use as a clinical tool and taking the most advantage of collected data by condensing data sets as little as possible.

However, it is not only the amount of data that is of relevance. The conceptualization of how useful measures can be derived from the collected data is of equal importance. Navigation data can be analyzed in many different ways. Arguably the most sophisticated solution is the use of pathfinding algorithms that are commonly applied in computer science, more specifically in video games. With such algorithms the environment is decomposed into a graph consisting of interconnected nodes. Each connection is assigned an edge path cost. Travelling along the nodes is represented by the summation of all edge path costs along the taken route. This allows the calculation of shortest routes and the cost of each path, potentially introducing a measurement of error which could be used to quantify navigation performance. However, the navigation task can potentially produce a large number of data points per second. Integrating this data set into a navigation graph (so-called navigation mesh) is not a straight forward task, especially for complex large-scale environments that involve several height levels (i.e. floors). The pathfinding setup and calculations would need to be adapted for the existing workflow. Before such expansion of the current system can be considered, the utility of the resulting data needs to be evaluated.

The data output of the VMT requires similar evaluation. Distance measures in three-dimensional space provide much more information than the conventional correct-incorrect dichotomy of many traditional neuropsychological test items. Consequently, a survey among potential clinical users could unveil useful solutions to display and process the system's collected data.

Additional concerns about the design and methodology of the conducted evaluations require further discussion. The combination of context-sensitive and process-specific aspects within the same application raises an interesting question. How does the individualized content of each task interfere with the task's standardization and comparison to normative data? Theoretically, the use of different target items and different context in memory assessments should have a profound influence on the task's outcome. Task context and target items are expected to be familiar and have relevance or emotional valence to the patient. Even though distances, angles, and task mechanics can be held constant across users, the task content will be different for each individualized application. Given that familiarity of the modeled environment and targets vary across patients and that extend of knowledge about the learning context influences memory performance (Smith, 1979), a standardization of the proposed tasks seems problematic.

In fact, it is unclear whether a standardization and collection of normative data is even necessary. As long as the individualized cognitive tasks possess predictive value and can provide information about the user's everyday performance to aid clinical decision-making, this controversy could be solved. However, such evaluation of predictive validity needs to be addressed in future studies. The outcome measures of the cognitive tasks need to be correlated with real-world functional tasks at different time intervals after the patient has been discharged from rehabilitation. Ylvisaker's (2003) approach to evaluating context-sensitive rehabilitation in a randomized controlled trial also addresses the described dilemma. A general evaluation of treatment efficacy can be achieved by using context as independent variable and comparing patients treated with non-individualized treatments against a group which receives context-sensitive treatment. Variation in such contextualized treatment as described by Ylvisaker would not be due to different tools and people involved in the delivery of the treatment (i.e. family members, teachers, coaches), but rather due to different VEs and target objects based on the same rehabilitation framework. The inclusion of functional performance after rehabilitation has finished would again serve as a main outcome measure and basis for assessing predictive validity.

The concept of context-dependent memory has been experimentally demonstrated in several studies (Anderson, 2000, pp. 279-280). Based on this body of evidence, the question arises whether this context effect can also be replicated with a combination of real and virtual context scenarios (e.g. acquisition in real environment, recall in virtual equivalent). This notion could lend strong support to assessments with high ecological validity, so that high-fidelity applications provide more accurate evidence for the patients' true abilities in environments that are relevant for everyday life. When taking these concerns into account, the comparison to traditional neuropsychological tests appears to be of lower priority, as the task's relationship to activities of daily living emerges as the more dominant concept. This situation could almost be seen as ironical, being the opposite of traditional validity concepts. As Shadish, Cook, and Campbell (2001) mention, ecological validity is the only characteristic that does not need to be present in order for an experiment to be considered valid. Internal and external validity on the other hand, are a requirement for the overall validity of an experiment (Shadish, et al., 2001). While it can be argued that test development as described in this dissertation and Shadish and colleagues' domain of experimental design conceptually differ, the controversy of the role of VR applications within the scope of validity evaluations remains. As previously noted, shifting focus from internal validity to comparisons of main outcome measures to functional tasks in a natural setting could shed light on the application's predictive validity. This is not to say that future evaluations of the proposed system should abandon the aim to reveal associations with well-established measures of cognition. It even has to be noted that the conducted experiments of this dissertation face several methodological limitations as outlined in this chapter. Due to these limitations it is unclear whether the absence of correlations of the proposed cognitive tasks with existing neuropsychological assessments are a result of the used methodology or are in fact the true outcome. Further evaluations are needed to address these limitations.

The fact that all evaluation studies have been conducted by one experimenter has to be noted as an additional limiting factor. Due to financial restraints study designs, assignment to experimental conditions as well as assessments/treatments have been carried out by the same person. Besides the experimenter being aware of the participants' assignment to conditions, participants were also aware of all experimental manipulations. As a consequence, non-blinded participants and experimenter can lead to biased results that need to be interpreted carefully. For all clinical trials (Chapter 3.2, 3.3, 3.4) the experimenter was also part of the therapeutic team for most of the participating patients. Even though the initial patient interview, information sheets and consent forms clearly indicated the strict separation of therapy and research participation, some patients might have not performed to their full abilities, depending on the cause of their deficits and their perspective beyond rehabilitation. For upcoming studies, data collection should be conducted by an experimenter who is unaware of the assignment of patients to different experimental conditions. A double-blind procedure seems not feasible as the experimental protocols and manipulations mostly are easy to distinguish for the participant and need to be explained in detail before establishing informed written consent.

A further limitation of the conducted trials is the deliberate choice of participants from a broad range of backgrounds. Healthy participants within a wide range of age, computer skills, and education were chosen for the navigation task evaluation (Chapter 3.1). Participants from the Christchurch community were haphazardly approached in public places throughout the city. The healthy participants were chosen due to limited availability of brain-injured patients and the prototypical nature of the VNT. Consequently, results of experiment one cannot be generalized towards clinical populations until further evaluation trials have been conducted. It can be assumed that a clinical population's performance on the VNT will be inferior to the outcomes of the current VNT trial. Though, without an additional trial this first experiment can merely serve as a comparison and guideline for future trials.

Clinical participants for Experiments 2-4 (Chapters 3.2, 3.3, 3.4) were recruited regardless of size and location of their brain injury. The tested clinical sample was comprised of patients across a wide range of age, computer experience, education, and cognitive abilities. These liberal recruitment strategies were employed to expose the developed applications to a heterogeneous set of users for comprehensive collection of user feedback. The proposed rehabilitation framework is designed to be applicable to a wide range of brain injuries that cause deficits of memory or spatial abilities. Also, the recruitment criteria present a balance between available resources (i.e. time, patients, funding, participating hospitals and clinicians) and the ability to draw meaningful conclusions from the trial. A well-defined homogeneous patient group requires a large pool of patients to choose from, preferably from several large hospitals over an extended period of time. If time or patient availability is restricted, the inclusion criteria can be broadened in order to collect sufficient data points for meaningful comparisons. Alternatively, few patients with similar lesion characteristics can be recruited for multiple testing sessions and extended exposure. However, neurological rehabilitation can be unpredictable with patients being transferred between hospitals, health care plans not paying for extended rehabilitation, seizures and headaches preventing test sessions and patients often lacking motivation or willingness to give informed consent. Hence, less restrictive inclusion criteria were chosen to avoid the risk of only recruiting very few patients or no patients at all in the limited timeframe of the clinical trials in Schaufling, Germany. The chosen recruiting strategy allowed for user feedback from a wide range of patients. However, the heterogeneous group of patients must be seen as a limitation in the context of the conducted validity analysis. The obtained effect sizes were small, in part due to the large variability of the data sets. With such low statistical power the employed analyses might have not been able to detect significant differences between the experimental conditions. Therefore, future trials need to consider the recruitment of homogeneous samples with well-defined inclusion and exclusion criteria. For clinical samples the recruitment of patients with distinct brain lesions (e.g. damage to the right posterior parietal cortex) seems to be a promising approach for reducing variability of collected data. The existing exploratory studies of this dissertation can then be used for conservative effect size estimation (i.e. assuming that following studies show less variability) to calculate sample sizes for adequate statistical power in upcoming clinical trials.

For future single-case trials the experimental protocol has to be extended to include enough training sessions to warrant a statistical time-series analysis (e.g. trend estimation). The data set for each session would optimally consist of information collected within the VE and additional well-established measures of cognitive abilities. Testing activities of daily living for each session can be a useful alternative. However, the availability of validated alternate versions of neuropsychological tests may become an issue for such protocol. The learning effects of repeated testing need to be taken into account, regardless of whether functional activities or pen and paper tests are used. Further, the capacity and daily fluctuations of the patients' attentional resources are limiting factors for extended testing/training sessions. Depending on the patient's therapy schedule and specific brain lesion, 30 to 60 minutes of cognitive tests and trainings were acceptable in most of the conducted sessions. With such limited time, tests and training tasks need to be chosen and balanced carefully to avoid introducing biased results due to the patient's exhaustion.

During the evaluation of the VMT it became apparent that the alternation of target items between trials caused interference for several participants. The occurrence and location of items in previous trials was reported to be interfering with the current task. The number of relevant target items depends on the task that the patient is expected to train. For example, preparing a breakfast involves a limited number of items so that for a comprehensive training some targets will have to be repeated over time. A future study needs to assess the influence of target pool size, target repetitions, familiarity with targets, and the interaction of those factors with the user's mnemonic strategies on task outcome.

Lastly, the motivation of users is expected to be high due to the realism and relevance of the VE and cognitive tasks. It can be argued that the use of game elements as part of the framework could potentially enhance the long-term motivation of users to continuously practice cognitive skills. Features like character progression, story-driven events, game scores and other game mechanics can provide means for users to train cognitive skills while having fun at the same time. However, the inclusion of game mechanics can require the implementation of abstract concepts which in turn have to be shown to transfer to the related real-world concepts. For example, Dunwell, Christmas and de Freitas (2011) describe the serious game "Code of Everand" in which children have to watch for incoming monsters from left and right before crossing "spirit channels". This serves as a metaphor for street crossing behavior in real-world scenarios. However, during the design and evaluation of the application it has to be taken into account that the user has to make the connection between metaphor and actual target behavior. Such requirement of abstraction was purposely avoided during this thesis by using realistic scenarios and cognitive tasks. Transfer of trained behavior to real-world situations has not received much support in clinical trials to date (Ylvisaker, 2003) and thus, the individualized simulation approach has been chosen over the development of a serious game.

4.3.4 Expansion of current framework

The proposed set of cognitive tasks is currently in a prototypical state and consists of individual VEs and a set of cognitive tasks. A fully functional VE can be developed in a matter of one to two days using the described workflow. However, in-depth knowledge of the 3D modeling process and the integration process within Unity are necessary to publish an application for a clinician to use. While the VMT has been developed by using extensive feedback from patients and clinicians, the navigation task was the first prototype developed for this dissertation. Its implementation and setup are not yet as user-friendly and require further development effort in order to be used effectively. Navigation targets and instructions need to be setup intuitively to compliment the rest of the proposed workflow. Specifically, Unity's interface needs to be extended through editor scripting in order to provide a user-friendly interface for setting up the navigation task's parameters and targets.

In the future, a comprehensive framework that extends across several cognitive domains is planned. Further cognitive tasks have to be developed and evaluated to give patients the opportunity to train a wide variety of cognitive abilities in relevant VEs depending on their individual needs. The context-sensitive approach that has been introduced in this dissertation will eventually be combined with domain-specific training tasks for prospective memory, problem-solving, cognitive flexibility, visual attention and several other sub-types of attentional processes (Zomeren & Brouwer, 1994). With the emergence of more effective tools for capturing and modeling 3D environments, the remaining barriers for the widespread use of individualized rehabilitation tools will be addressed. In the meantime, next steps towards making this rehabilitation approach a reality include further validation studies and a randomized clinical trial. The advantages of individualized VR rehabilitation over traditional rehabilitation and the use of generic VR scenarios need to be demonstrated repeatedly before widespread clinical use becomes a possibility. While the benefit of this individualized framework above and beyond current practices has not been demonstrated yet, the value of VR-based individualized cognitive task for predicting everyday task performance in the unique context of a patient's everyday life needs to be considered. Depending on the complexity of this context and the demands placed on the patient, individualized VR-tasks may provide valuable information to the clinician. In some cases a standardized pen and paper assessment might be sufficient; in other cases a hands-on functional task might provide more information. As long as the use of this individualized approach provides additional valuable information for the rehabilitative process of some patients, it is worth considering this framework as an additional tool amongst other assessment and training choices. It is for upcoming experimental trials to find out in which situations patients can profit the most from this approach.

An important outcome of this thesis work is the anecdotal finding that five patients gained insights in their cognitive deficits while using the VMT (chapter 3.2). This can potentially be an important factor for using realistic VEs for cognitive rehabilitation. However, deficit denial was not assessed as part of the experimental trials and was only discovered incidentally during patient debriefing and the reaction of the patients during the test session. Future clinical trials need to specifically address this outcome and systematically evaluate the effect of high ecological validity on deficit awareness. This is important for ethical implications of patients coming to terms with the drastic impact that brain-injuries can have on their daily life. Also, deficit awareness appears to be critically important for the success of cognitive rehabilitation (Ownsworth & Clare, 2006). Consequently, this factor should be considered to be one of the main goals for future trials and can be seen as a potential key factor in using VR technology for rehabilitation.

Lastly, it needs to be stated that the intention of this dissertation was solely to develop and evaluate a prototypical set of cognitive tasks embedded in individual VEs. This initial development effort is believed to be a step towards promoting and fostering further research at the intersection of context-sensitive cognitive rehabilitation and VR. The resulting motivation, deficit awareness and feedback of users are important outcomes to build upon in future trials. During the course of this work aspects of cost-effectiveness have repeatedly been pointed out by fellow researchers as a critical factor for this rehabilitation approach. Questions of development costs, distribution of the environments, and profitability arose. However, considering the speed at which technologies evolve in the 21st century, many of the technical limitations of the presented framework will be obsolete in the near future. Instead of worrying about the framework's integration in current health care budgets, the theoretical foundations and the long-term benefits of individualized VR rehabilitation have to be shown first.

5 BIBLIOGRAPHY

Adams, K. M. (2003). Realising the Potential of Cognitive Rehabilitation for the Brain-Injured: Next Steps. *Brain Impairment, 4*(1), 31-35.

Anderson, J. R. (2000). *Learning and Memory* (2nd ed.). New York: John Wiley.

APA, NCME, & AERA. (1999). The Standards for Educational and Psychological Testing. Washington D.C.

Autodesk Inc. Autodesk Design Review Retrieved 10/16, 2011, from http://usa.autodesk.com/design-review/

Ben-Yishay, Y., Piasetsky, E. B., & Rattok, J. (1987). A systematic method for ameliorating disorders in basic attention. In A. L. B. Manfred J. Meier, Leonard Diller (Ed.), *Neuropsychological Rehabilitation* (pp. 475). New York: The Guilford Press.

betanet. (2011). Anschlussheilbehandlung [Follow-up treatment] Retrieved 10/08, 2011, from http://www.betanet.de/betanet/soziales_recht/Anschlussheilbehandlung-17.html

Bowman, D. A., Datey, A., Ryu, Y. S., Farooq, U., & Vasnaik, O. (2002). *Empirical comparison of human behavior and performance with different display devices for VEs.* Paper presented at the Human Factors and Ergonomics Society Annual Meeting.

Bowman, D. A., Koller, D., & Hodges, L. F. (1998). A Methodology for the Evaluation of Travel Techniques for Immersive virtual environments. *Virtual Reality, 3*, 120-131.

Brickenkamp, R. (1981). *Test d2: Aufmerksamkeits-Belastungstest, Handanweisung 7th Ed. [d2 Test of Attention, Manual 7th Ed.]*. Göttingen: Verlag für Psychologie Dr. C. J. Hogrefe.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix. *Psychological Bulletin, 56*(2), 81-105.

Cicerone, K. D., Dahlberg, C., Kalmar, K., Langenbahn, D. M., Malec, J. F., Bergquist, T. F., Felicetti, T., Giacino, J. T., Harley, P. J., Harrington, D. E., Herzog, J., Kneipp, S., Laatsch, L., & Morse, P. A. (2000). Evidence-Based Cognitive Rehabilitation: Recommendations for Clinical Practice. *Arch Phys Med Rehabil, 81*, 1596-1615.

Cicerone, K. D., Dahlberg, C., Malec, J. F., Langenbahn, D. M., Felicetti, T., Kneipp, S., Ellmo, W., Kalmar, K., Giacino, J. T., Harlez, P., Laatsch, L., Morse, P., & Catanese, J. (2005). Evidence-Based Cognitive Rehabilitation: Updated Review of the Literature From 1998 Through 2002. *Arch Phys Med Rehabil, 86*, 1596-1615.

Cohen, J. (1987). *Statistical power analysis for the behavioral sciences*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Conroy, B. E., DeJong, G., & Horn, S. D. (2009). Hospital-Based Stroke Rehabilitation in the United States. *Top Stroke Rehabil, 16*(1), 34-43.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-Experimentation: Design and Analysis for Field Settings*. Chicago: Rand McNally.

Curtis, L. E., Siegel, A. W., & Furlong, N. E. (1981). Developmental Differences in Cognitive Mapping: Configurational nowledge of Familiar Large-Scale Environments. *Jorunal of Experimental Child Psychology, 31*, 456-469.

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

Dunwell, I., Christmas, S., & Freitas, S. d. (2011). Code of Everand - Final Evaluation Report: Serious Games Institute.

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Kit of factor-referenced cognitive tests*. NJ, Princeton: Educational Testing Service.

Freedman, S. A., Hoffman, H. G., A., G.-P., Weiss, P. L., Avitzour, S., & Josman, N. (2010). Prolonged Exposure and Virtual Reality-Enhanved Imaginal Exposure for PTSD following a Terrorist Bulldozer Attack: A Case Study. *CyberPsychology and Behavior, 13*(1), 95-101.

Furness, T. I. (1992). Towards tightly-coupled human interfaces. In E. Rae, G. Richard, v. D. Andries & V. John (Eds.), *Frontiers of Human-Centered Computing, Online Communities and* virtual environments (pp. 80-98). London: Springer Verlag.

Garcia-Palacios, A., Hoffman, H., Carlin, A., Furness III, T. A., & Botella, C. (2002). Virtual reality in the treatment of spider phobia: a controlled study. *Behaviour Research and Therapy,* 40, 983-993.

Geusgens, C., Winkens, I., Heugten, C. v., Jolles, J., & Heuvel, W. v. d. (2007). The occurence and measurement of transfer in cognitive rehabilitation: A critical review. *Journal of Rehabilitation Medicine, 39*(6), 425-439.

Goldberg, H. (2006). Computer/Internet Experience Skills Questionnaire for: Internet Diabetes Trial at Harborview 2006. Retrieved 10/14, 2011, from http://www.hetinitiative.org/grantee_instruments/UWashHart_Baseline%20Computer%20Skills%20Survey.pdf

Golz, D., Huchler, S., Jörg, A., & Küst, J. (2004). Beurteilung der Fahreignung. *Zeitschrift für Neuropsychologie, 15*(3), 157-167.

Gorini, A., & Riva, G. (2008). Virtual Reality in anxiety disorder: the past and the future. *Expert Reviews of Neurotherapeutics, 8*(2), 215.

Hegarty, M., Richardson, A. E., Montello, D. R., Lovelace, K., & Subbiah, I. (2002). Development of a self-report measure of environmental spatial ability. *Intelligence, 30*, 425-447.

Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking abilities. *Intelligence, 32*, 175-191.

Henry, D., & Furness T. (1993). Spatial perception in virtual environments: Evaluating an architectural application. *Virtual Reality Annual International Symposium*, 33-40.

Izadi, S., Kim, D., Hilliges, O., Molyneaux, D., Newcombe, R., Kohli, P., Shotton, J., Hodges, S., Freeman, D., Davison, A., & Fitzgibbon, A. (2011). *KinectFusion: Real-time 3D Reconstruction and Interaction Using a Moving Depth Camera*. Paper presented at the ACM Symposium on User Interface Software and Technology, Santa Barbara.

Johnston, S. C., Mendis, S., & Mathers, C. D. (2009). Global variation in stroke burden and mortality: estimates from monitoring, surveillance, and modelling. *The Lancet Neurology, 8*, 345-354.

Kaiser Family Foundation. (2011a). Health Care Spending in the United States and Selected OECD Countries, April 2011 Retrieved 10/08/2011, 2011

Kaiser Family Foundation. (2011b). Total Health Expenditure per Capita, U.S. and Selected Countries, 2008; OECD Data.

Katz, N., Ring, H., Naveh, Y., Kizony, R., Feintuch, U., & Weiss, P. L. (2005). Interactive virtual environment training for safe street crossing of right hemisphere stroke patients with Unilateral Spatial Neglect. *Disability and Rehabilitation, 27*(20), 1235-1243.

Kennedy, R. S., Drexler, J. M., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. *International Journal of Aviation Psychology, 3*(3), 203-220.

Kessels, R. P. C., Zandvoort, M. J. E. v., Postma, A., Kappelle, L. J., & Haan, E. H. F. d. (2000). The Corsi Block-Tapping Task: Standardization and Normative Data. *Applied Neuropsychology, 7*(4).

King, J. A., Burgess, N., Hartley, T., Vargha-Khadem, F., & O'Keefe, J. (2002). Human Hippocampus and Viewpoint Dependence in Spatial Memory. *Hippocampus, 12*, 811-820.

Koenig, S., Crucian, G., Dalrymple-Alford, J., & Dünser, A. (2010). *Assessing navigation in real and* virtual environment*s: A validation study*. Paper presented at the 8th Intl Conf. on Disability, Virtual Reality and Assoc. Technologies, Viña del Mar/Valparaíso, Chile.

Koenig, S., Crucian, G., Dalrymple-Alford, J., & Dünser, A. (2011). Assessing navigation in real and virtual environments: A validation study Retrieved 10/13, 2011, from http://www.reference-global.com/doi/abs/10.1515/IJDHD.2011.050

Koenig, S., Crucian, G. P., Dünser, A., Bartneck, C., & Dalrymple-Alford, J. C. (2011a). *Development of a spatial memory task in realistic* virtual environment*s*. Paper presented at the Virtual Reality International Conference, Laval, France. http://www.laval-virtual.org/2011/papers_2011/67.pdf

Koenig, S., Crucian, G. P., Dünser, A., Bartneck, C., & Dalrymple-Alford, J. C. (2011b). Validity evaluation of a spatial memory task in virtual environments. *International Journal of Design and Innovation Research, 6*(1).

Koenig, S., Dünser, A., Bartneck, C., Dalrymple-Alford, J. C., & Crucian, G. P. (2011). *Development of* virtual environment*s for patient-centered rehabilitation*. Paper presented at the International Conference on Virtual Rehabilitation, Zurich, Switzerland. http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5971838

Koh, G., Wiegand, v. T. E., Garnett, R. L., Durlach, N. I., & Shinn-Cunningham, B. (1999). Use of virtual environments for acquiring configurational knowledge about specific real-world spaces: I. Preliminary experiment. *Presence, 8*(6), 632-656.

Kozhevnikov, K., & Hegarty, M. (2001). A dissociation between object-manipulation spatial ability and spatial orientation ability. *Memory and Cognition, 29*, 745-756.

Lange, B., Flynn, S., Chang, C.-Y., Ahmed, A., Geng, Y., Utsav, K., Xu, M., Seok, D., Cheng, S., & Rizzo, A. (2010). Development of an interactive rehabilitation game using the Nintendo (WiiFit) Balance Board for people with neurological injury. *Topics in Stroke Rehabilitation, 15*(5), 345-352.

Lange, B., Rizzo, A., Chang, C.-Y., Suma, E., & Bolas, M. (2011). *Markerless Full Body Tracking: Depth-Sensing Technology within* virtual environment*s*. Paper presented at the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC), Orlando.

Lewis, T. M., Aggarwal, R., Rajaretnam, N., Grantcharov, T. P., & Darzi, A. (2011). Training in surgical oncology - The role of virtual reality simulation. *Surgical Oncology, 20*, 134-139.

Livingstone, S. A., & Skelton, R. W. (2007). Virtual environment navigation tasks and the assessment of cognitive deficits in individuals with brain injury. *Behavioural Brain Research, 185*, 21-31.

Mercier, L., Therese, A., Hebert, R., Rochette, A., & Dubois, M.-F. (2001). Impact of Motor, Cognitive, and Perceptual Disorders on Ability to Perform Activitires of Daily Living After Stroke. *Stroke, 32*, 2602-2608.

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups design. *Psychological Methods, 7*, 105-125.

Nadolne, M. J., & Stringer, A. Y. (2001). Ecologic validity in neuropsychological assessment: Predicition of wayfinding. *Journal of the International Neuropsychological Society, 7*, 675-682.

New Zealand Guidelines Group. (2006). *Traumatic Brain Injury: Diagnosis, Acute Management and Rehabilitation*. Wellington: New Zealand Guidelines Group for ACC.

Nickerson, R. S. (2000). Null hypothesis significance testing: A review of an old and continuing controversy. *Psychological Methods, 5*(2), 241-301.

Olbrich, R. (1996). Computer based psychiatric rehabilitation: current activities in Germany. *Eur Psychiatry, 1996*(11, Suppl. 2.), 60-65.

Osterrieth, P. A. (1944). The test of copying a complex figure: A contribution to the study of perception and memory (Filetest de copie d'une figure complex: Contribution a l'etude de la perception et de la memoire). *Archives de Psychologie, 30*, 286-356.

Owen, A. M., Hampshire, A., Grahn, J. A., Stenton, R., Dajani, S., Burns, A. S., Howard, R. J., & Ballard, C. G. (2010). Putting brain training to the test. *Nature, 465*, 775-778.

Ownsworth, T., & Clare, L. (2006). The association between awareness deficits and rehabilitation outcome following acquired brain injury. *Clinical Psychology Review, 26*(6), 783-795.

Parsons, T. D., Kenny, P., Ntuen, C. A., Pataki, C. S., Pato, M. T., Rizzo, A. A., St-George, C., & Sugar, J. (2008). Objective Structured Clinical Interview Training using a Virtual Human Patient. *Medicine Meets Virtual Reality, 16*, 357-362.

Patel, M. D., Coshall, C., Rudd, A. G., & Wolfe, C. D. A. (2002). Cognitive Impairment after Stroke: Clinical Determinants and Its Associations with Long-Term Stroke Outcomes. *Journal of the American Geriatrics Society, 50*(4), 700-706.

Population Reference Bureau. (2010). *2010 World Population Data Sheet*. Washington D.C.: Population Reference Bureau.

Prensky, M. (2002). The motivation of gameplay. *On the Horizon, 10*(1), 5-11.

Price, M., Anderson, P., & Rothbaum, B. (2008). Virtual reality as treatment for fear of flying - a review of recent research. *The International Journal of Behavioral Consultation and Therapy, 4*(4), 340-347.

Rao, R., Jackson, S., & Howard, R. (1999). Neuropsychological Impairment in Stroke, Carotid Stenosis, and Peripheral Vascular Disease: A Comparison With Healthy Community Residents. *Stroke, 30*, 2167-2173.

Reynish, W. (2008). The Evolution of Blenders User Interface.

Richardson, A. E., Montello, D. R., & Hegarty, M. (1997). Spatial knowledge acquisition from maps and from navigation in real and virtual environments. *Memory and Cognition, 27*(4), 741-750.

Riva, G., Gaggioli, A., Davide Borra, Gagliati, A., DonVito, G., Gasco, S., Cipresso, P., Pallavicini, F., Raspelli, S., Grassi, A., & Vigna, C. (2011). NeuroVR 2.0 Retrieved 10/16, 2011, from http://www.neurovr2.org/

Rizzo, A., Difede, J., Rothbaum, B. O., Johnston, S., McLay, R. N., Reger, G., Rahm, G., Parsons, T., Graap, K., & Pair, J. (2009). Virtual Reality PTSD Exposure Therapy Results with Active Duty OIF/OEF Combatants. *Medicine Meets Virtual Reality, 17*, 277-282.

Rizzo, A., Difede, J., Rothbaum, B. O., Reger, G., Spitalnick, J., Cukor, J., & Mclay, R. (2010). Development and early evaluation of the Virtual Iraq/Afghanistan exposure therapy system for combat-related PTSD. *Annals of The New York Academy of Sciences, 1208*, 114-125.

Rizzo, A., Graap, K., McLay, R. N., Rothbaum, B. O., Difede, J., Reger, G. M., Holloway, K., Gahm, G. A., Johnston, S., Spitalnick, J., Bordnick, P., Newman, B., & Parsons,

T. D. (2009). Clinical Results from a Virtual Iraq PTSD Exposure Therapy Application. *Frontiers in Neuroengineering: Conference Abstract: Annual CyberTherapy and CyberPsychology 2009 conference*, 8-15. doi: 10.3389/conf.neuro.14.2009.06.088

Rizzo, A., & Kim, G. J. (2005). A SWOT Analysis of the Field of Virtual Reality Rehabilitation and Therapy. *Presence, 14*(2), 119-146.

Rizzo, A., Schultheis, M., Kerns, K. A., & Mateer, C. (2004). Analysis of assets for virtual reality applications in neuropsychology. *Neuropsychological Rehabilitation, 14*(1/2), 207-239.

Rose, F. D., Brooks, B. M., & Rizzo, A. A. (2005). Virtual reality in brain damage rehabilitation: Review. *CyberPsychology and Behavior, 8*(3), 241-263.

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

Sangani, S., Weiss, P. L., Kizony, R., Koenig, S. T., Levin, M. F., & Fung, J. (2012). *Development of a complex ecological* virtual environment. Paper presented at the Proceedings of the 9th International Conference on Disability, Virtual Reality and Associated Technologies, Laval, France.

Santos, B. S., Dias, P., Pimentel, A., Baggerman, J.-W., Ferreira, C., Silva, S., & Madeira, J. (2009). Head-mounted display versus desktop for 3D navigation in virtual reality: a user study. *Multimedia Tools and Applications, 41*, 161-181.

Schretlen, D., & Shapiro, A. (2003). A quantitative review of the effects of traumatic brain injury on cognitive functioning. *International Review of Psychiatry, 15*, 341-349.

Schuirmann, D. (1987). A comparison of the Two One-Sided Tests Procedure and the Power Approach for assessing the equivalence of average bioavailability. *Journal of Pharmacokinetics and Pharmacodynamics, 15*(6), 657-680.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*: Wadsworth Publishing.

Shrager, Y., Bayley, P. J., Bontempi, B., Hopkin, R. O., & Squire, L. R. (2006). Spatial memory and the human hippocampus. *Proceedings of the National Academy of Sciences, 104*(8), 2961-2966.

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B - Biological Sciences, 364*, 3549-3557.

Sloan, S., Winkler, D., & Callaway, L. (2004). Community Integration Following Severe Traumatic Brain Injury: Outcomes and Best Practice. *Brain Impairment, 5*(1), 12-29.

Smith, S. M. (1979). Remembering in and out of context. *Journal of Experimental Psychology: Leanring, Memory, and Cognition, 5*(5), 460-471.

Sohlberg, M. M., & Mateer, C. A. (1989). *Introduction to cognitive rehabilitation: theory and practice*. New York: Guilford Press.

Sohlberg, M. M., & Mateer, C. A. (Eds.). (2001). *Cognitive rehabilitation: An integrative neuropsychological approach*. New York: Guilford Press.

Stanney, K. (Ed.). (2002). *Handbook of* virtual environment*s: Design, Implementation, and Applications*. Mahwah, NJ: Lawrence Erlbaum Associates.

Stanney, K. M., Kingdon, K. S., Graeber, D., & Kennedy, R. S. (2002). Human Performance in Immersive virtual environments: Effects of Duration, User Control, and Scene Complexity. *Human Performance, 15*(4), 339-366.

Stanney, K. M., Mollaghasemi, M., Reeves, L., Breaux, R., & Graeber, D. A. (2003). Usability engineering of virtual environments (VEs): identifying multiple criteria that drive effective VE system design. *International Journal of Human-Computer Studies, 58*, 447-481.

Sturm, W., Willmes, K., Orgass, B., & Hartje, W. (1997). Do Specific Attention Deficits Need Specific Training? *Neuropsychological Rehabilitation, 7*(2), 81-103.

Trenholme, D., & Smith, S. P. (2008). Computer game engines for developing first-person virtual environments. *Virtual Reality, 12*, 181-187.

Tryon, W. W. (2001). Evaluating statistical difference, equivalence, and indeterminancy using inferential confidence intervals: An integrated alternative method of conducting null hypothesis statistical tests. *Psychological Methods, 6*(4), 371-386.

Tryon, W. W., & Lewis, C. (2008). An inferential confidence interval method of establishing equivalence that corrects Tryon's (2001) reduction factor. *Psychological Methods, 13*(3), 272-277.

Tucha, L., Tucha, O., Laufkötter, R., Walitza, S., Klein, H. E., & Lange, K. W. (2008). Neuropsychological assessment of attention in adults with different subtypes of attention-deficit/hyperactivity disorder. *Journal of Neural Transmission, 115*, 269-278.

Vakhnina, N. V., Nikitina, L. Y., Parfenov, V. A., & Yakhno, N. N. (2009). Post-Stroke Cognitive Impairments. *Neuroscience and Behavioral Physiology, 39*(8), 719-724.

van Asselen, M., Kessels, R. P. C., Kappelle, L. J., Neggers, S. F. W., Frijns, C. J. M., & Postma, A. (2006). Neural correlates of human wayfinding in stroke patients. *Brain Research, 1067*(1), 229-238.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills, 47*, 599-604.

Vincelli, F., Anolli, L., Bouchard, S., Wiederhold, B. K., Zurloni, V., & Riva, G. (2003). Experimental Cognitive Therapy in the Treatment of Panic Disorders with Agoraphobia: A Controlled Study. *CyberPsychology and Behavior, 6*(3), 321-328.

Virvou, M., & Katsionis, G. (2008). On the usability and likeability of virtual reality games for education: The case of VR-ENGAGE. *Computers and Education, 50*, 154-178.

Wade, D. T. (2000). Personal context as a focus for rehabilitation. *Clinical Rehabilitation, 14*, 115-118. doi: 10.1191/026921500672636483

Wade, D. T. (2001). Social context as a focus for rehabilitation. *Clinical Rehabilitation, 15*, 459-461. doi: 10.1191/026921501680425171

Walker, W. C., & Treven, P. C. (2007). Motor impairment after severe traumatic brain injury: A longitudinal multicenter study. *Journal of Rehabilitation Research and Development, 44*(7), 975-982.

Waller, D. (2000). Individual Differences in Spatial Learning From Computer-Simulated Environments. *Journal of Experimental Psychology: Applied, 6*(4), 307-321.

Wechsler, D. (1945/1997). *Wechsler Memory Scale: Third edition (WMS-III)*. San Antonio: Harcourt Assessment.

Weiand, C. (2006). *Neuropsychologische Behandlungsmethoden im Vergleich - Eine randomisierte klinische Studie [Comparing neuropsychological treatment methods - A randomized clinical trial].* PhD dissertation, University of Konstanz, Konstanz.

Werner, P., Rabinowitz, S., Klinger, E., Korczyn, A., & Josman, N. (2009). Use of the Virtual Action Planning Supermarket for the Diagnosis of Mild Cognitive Impairment. *Dement Geriatr Cogn Disord, 27*, 301-309.

Westerberg, H., Jacobaeus, H., Hirvikoski, T., Clevberger, P., Östensson, M.-L., Bartfai, A., & Klingberg, T. (2007). Computerized working memory training after stroke – A pilot study. *Brain Injury, 21*(1), 21-29.

Wilson, B. A. (2000). Compensating for Cognitive Deficits Following Brain Injury. *Neuropsychology Review, 10*(4), 233-243.

Witmer, B. G., Bailey, J. H., Knerr, B. W., & Parsons, K. C. (1996). Virtual Spaces and real world places: transfer of route knowledge. *Int J Hum Comput Stud, 45*, 413-428.

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

World Health Organization. (2001). International Classification of Functioning, Disability and Health (ICF): World Health Organization,.

Ylvisaker, M. (2003). Context-sensitive Cognitive Rehabilitation after Brain Injury: Theory and Practice. *Brain Impairment, 4*(1), 1-16.

Ylvisaker, M. (2006). Self-Coaching: A Context-Sensitive, Person-Centred Approach to Social Communication After Traumatic Brain Injury. *Brain Impairment, 7*(3), 246-258.

Zomeren, A. H. v., & Brouwer, W. H. (1994). *Clinical Neuropsychology of Attention*. New York: Oxford University Press.