Validity evaluation of a spatial memory task in virtual environments

Sebastian T Koenig[1,2], Gregory P Crucian[3], Andreas Dünser[1], Christoph Bartneck[1], John C Dalrymple-Alford[2]

[1] Human Interface Technology Laboratory New Zealand
Private Bag 4800
8140 Christchurch, New Zealand
sebastian.koenig@canterbury.ac.nz

[2] Department of Psychology, University of Canterbury
Private Bag 4800
8140 Christchurch, New Zealand

[2] Department of Behavioral Health : William Beaumont Army Medical Center
5005 N Piedras St.
79920-5001 El Paso, TX, USA

ABSTRACT. *A broad range of neuropsychological tests and trainings are used to assess cognitive function during rehabilitation after brain injury. However, given the everyday problems that patients struggle with, questions of ecological validity arise. For example, abstract tasks have little meaning to patients and decrease motivation to practice skills frequently. This study proposes a spatial memory task with high ecological validity that can be integrated into any virtual environment. Environments and target objects can be individually designed for each user to provide a relevant context and high motivation for patients with cognitive deficits. The task has been evaluated with both brain-injured and healthy individuals. 45 participants completed the virtual task and a battery of neuropsychological tests. Strong associations have been found with established tests of visual short-term memory. Due to high variability of test scores, no significant relationship with tests of spatial abilities has been established. High ecological validity of the spatial memory task caused five patients to show awareness of their cognitive deficits. Clinical implications and future task development are discussed.*

KEYWORDS: virtual environment; neuropsychology; rehabilitation; ecological validity; brain injury; spatial memory, validity evaluation.

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 and RehaCom). Modern commercial cognitive tasks, brain teasers, and 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 virtual environments. The Virtual Memory Task (VMT) was designed to provide a clinical tool which has several advantages over traditional cognitive tasks:

- higher ecological validity by using personalized, realistic virtual environments
- 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 task is a software module (so-called asset) for the Unity game engine that is easy to set up by placing it into an existing virtual environment via drag&drop. The approach of a modular task enables the researcher and clinician to create an assessment or training in any virtual environment which is relevant to the patient. This approach is limited by the resources it takes to create a virtual model of the individual environment. However, with modern game engines, high productivity tools and an optimized workflow, such obstacles no longer prevent any brain-injured patient from receiving individualized rehabilitation tools. The VMT's rationale has been inspired by experiments of King and colleagues (2002) and Shrager and colleagues (2006) in which virtual reality tasks were used to assess spatial memory of brain-injured individuals. 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. The VMT was designed to involve a combination of short-term memory and perspective taking skills. Hence, correlations with neuropsychological tests measuring those constructs are expected. Integration into clinical context, usability, task development, and task validation are discussed in detail.

2. Methods

2.1. *Participants*

45 participants (22 male - 23 female) at the Neurological Department of the Asklepios Klinik 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 maintain performance for at least 30 minutes. Computer experience was not required for participation. All patients were able to give informed written consent.

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.

2.3. *Materials*

2.3.1. *Pen and Paper Tests*

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 block span and digit span of the Wechsler Memory Scale III [forward and backward; WMS 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 Object Perspective Taking Test 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 Mental Rotations Test consists of twenty items which require a comparison between a target object and four test objects. All objects are random three-dimensional line drawings. Two of the four test objects are rotated, but identical versions of the target object and need to be identified by the participant. The test's score is calculated by dividing the number of correctly identified test objects by the number of attempted test objects.

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 and omission and false positive errors. The total number of processed targets minus the number of errors is used as the test's score.

2.3.2. *Virtual Memory Task (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 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 virtual environment followed the procedures as described by Koenig and colleagues [2011].

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 (Fig. 1).

Prior to the first task the participant was shown an overview of the surrounding virtual environment 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 move 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 (Fig. 2). 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 continuous motion towards a new location in the virtual environment with the user's virtual field of view always centered on the virtual table.

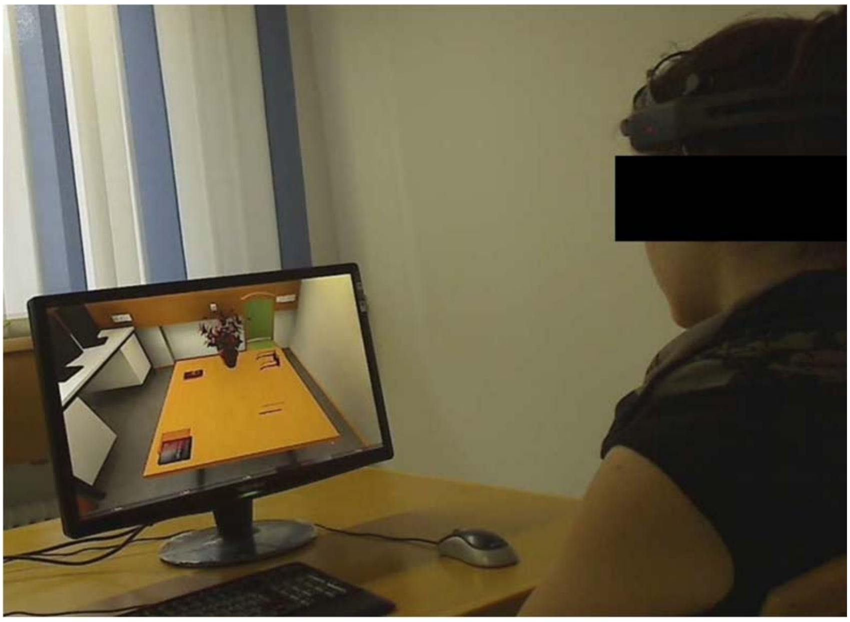

Figure 1. *Participant completing the VMT and receiving feedback about distance errors (red, transparent targets)*

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 virtual environment, 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. Distance error and all target positions were saved as text files for each trial. Rotation of target items was not relevant for this experiment.

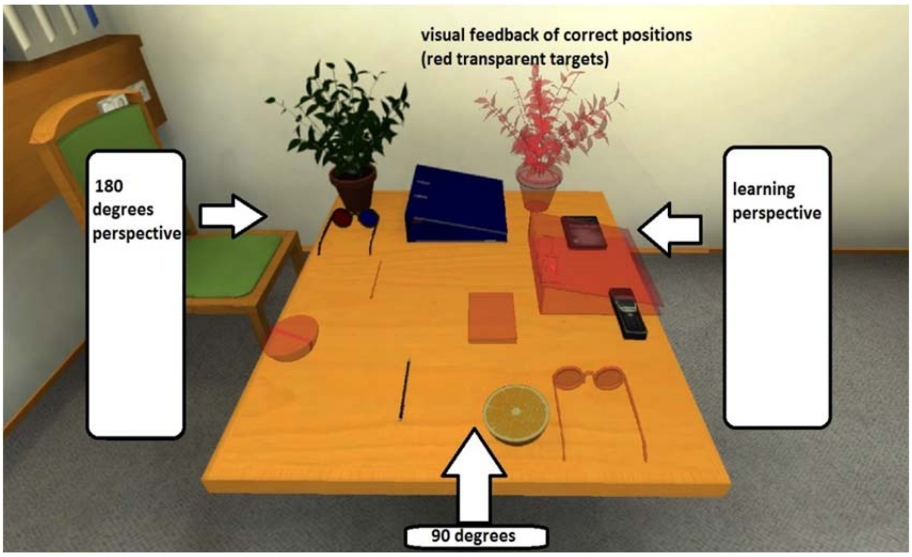

Figure 2. *Perspective changes for Virtual Memory Task (white rectangles represent user positions and are not part of the task)*

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 are beyond the scope of this paper and will be discussed elsewhere. 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.

2.5. *Statistical Analysis*

Statistical analyses were carried out using the software PASW 18 (IBM SPSS Inc.). 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. 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. Results

The participants' performance on the VMT was subject to two Friedman tests, analyzing each of the test's two factors separately –perspective change (Fig. 3) and number of target items (Fig. 4). 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 ($\alpha = 0.05/9 = 0.0055$). Significant differences were found for comparisons for 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 0.005 ($\alpha = 0.05/10$). 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 1.

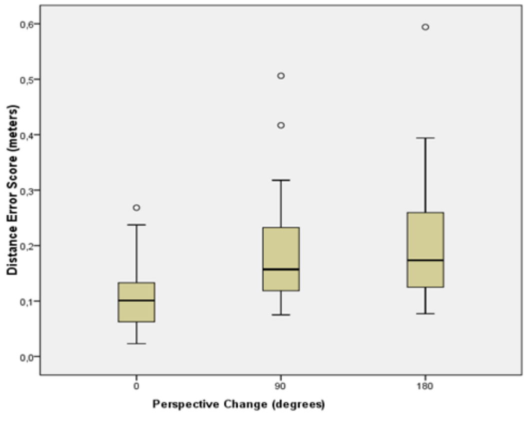

Figure 3. *Boxplot for distribution of Distance Error Score and Perspective Change*

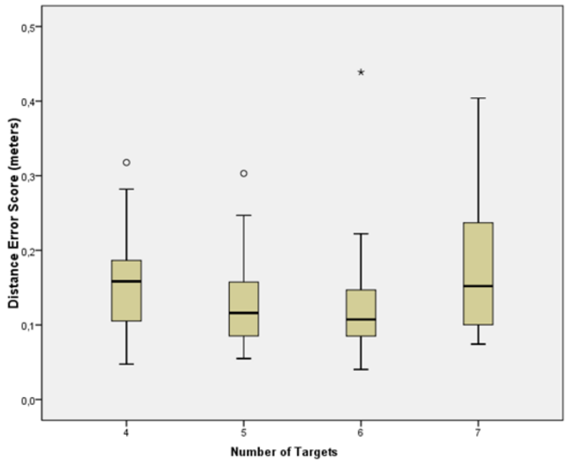

Figure 4. *Boxplot for distribution of Distance Error Score and Number of Targets*

Table 1. *Spearman's rank-order correlations of VMT distance error score with measures of memory and attention*

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

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. 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 task 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 virtual environment that is compatible with 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. This match between real and virtual test environment was supposed to give the participant a better sense of space during the 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 able to move the targets via instructions by the participant. This was done for one patient with no computer experience at all. For patients with more severe tremor, a modified USB-numeric keypad with large keys has been used to move targets onscreen.

The primary goal of this study was to demonstrate the test's convergent and discriminant validity in a clinical context. It was expected that test performance on the VMT shows strong significant correlations with measures of short-term memory and working memory. It was also expected that for trials, in which a perspective change is applied, a significant correlation with spatial abilities tests is evident. No significant correlation was expected between VMT performance and measures of attention.

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. No significant relationship has been shown between VMT error scores and tests of spatial abilities. 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 differ 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 (Fig. 5-A; i.e. ignoring a perspective change), distance between targets (Fig. 5-B; with correct layout of targets), total shift of the array of items (Fig. 5-C; e.g. when misinterpreting foreshortening of camera perspective), swapping target locations (Fig. 5-D), or location of single targets (Fig. 5-5; *"I forgot where it was"*).

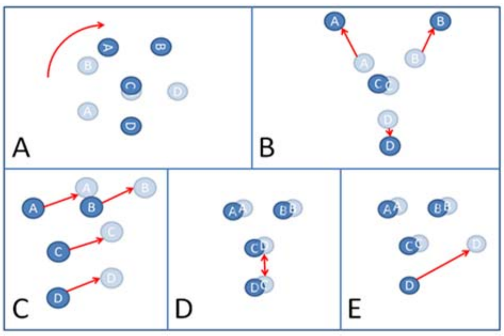

Figure 5. *Observed Error Types for VMT*

While developing and testing prototypes of sophisticated analysis tools, errors D and E (Fig. 5) 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 on a two-dimensional grid in the virtual environment, as deviations from the common midpoint of the set of targets. Figure six 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 is a Unity-based standalone tool which reads the text files that the VMT produces. It also visualizes 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.

Figure 6. *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 [2007], 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 virtual environments, it is necessary to use separate locations for virtual environment 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. This factor becomes more important with realistic virtual environments that allow the user to recognize their surroundings.

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 generally, participants committed larger errors for trials with higher number of target items and also for larger perspective changes. No significant differences were found 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 current study provided a first-hand experience of how participants act in a realistic, semi-familiar virtual environment. 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. An important finding of this study is the strong effect that the high ecological validity had on the patients' awareness of their cognitive deficits. Five cases were identified in which patients with mnestic deficits went through all well-established 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 common, 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 is not recommended at this stage and therapeutic potential of tasks with high ecological validity need to be extensively tested in future trials.

The development of the VMT task has been focused on creating a modular task which can be easily placed into any virtual environment. The workflow of creating a virtual environment has been refined to allow for a quick prototyping of virtual spaces in a matter of hours. Using Google SketchUp and the Google Warehouse to import freely available furniture and household items into the virtual environment so far has proven to be the quickest way for creating virtual spaces. 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 virtual environments and their long-term usage during rehabilitation must be subject to further investigation.

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 their 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 virtual environment (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 proven to be a task of high ecological validity which even made it possible for several highly-impaired individuals to realize their cognitive deficits. Usability and user feedback have been excellent throughout so that further trials and extended use of the application during cognitive rehabilitation seem justified. However, when using the VMT with patients with cognitive impairments, continuous support by experienced therapists is recommended to avoid frustration. High ecological validity and realism of the virtual task 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.

Acknowledgment

This study has been supported by AaronC of Deepwater3D.com, Asklepios Rehabilitation Clinic (Schaufling, Germany), and HIT Lab New Zealand.

References

Brickenkamp R., *Test d2 : Aufmerksamkeits-Belastungstest, Handanweisung 7th ed.* Göttigen: Verlag für Psychologie Dr. C. J. Hogrefe, 1981.

Goldberg, H. "Computer/Internet Experience Skills Questionnaire for: Internet Diabetes Trial at Harborview 2006. Accessed 2011 Jun 09. URL: http://www.hetinitiative.org/grantee_instruments/UWashHart_Baseline%20Computer%20Skills%20Survey.pdf

Hegarty M., Waller D., "A dissociation between mental rotation and perspective-taking spatial abilities," *Intelligence*, Vol. 32, 2004, pp. 175-91.

King J. A., Burgess N., Hartley T., Vargha-Khadem F., O'Keefe J., "Human hippocampus and viewpoint dependence in spatial memory", *Hippocampus*, Vol.12, 2002, pp. 811-820.

Koenig S. T., Dünser A., Bartneck C., Dalrymple-Alford J. C., Crucian G. P., „Development of virtual environments for patient centered rehabilitation", 2011, *Proceedings of the International Conference on Virtual Rehabilitation*, Zürich.

Osterrieth P. A., "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*, Vol. 30, 1944, pp. 286-356.

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

Shrager Y., Bayley P. J., Bontempi B., Hopkins R. O., Squire L. R., "Spatial memory and the human hippocampus", *Proceedings of the National Academy of Sciences*, Vol. 104, No. 8, 2007, pp. 2961-2966.

Sohlberg M. M., Mateer C. A., *Cognitive rehabilitation: An integrative neuropsychological approach.* New York: Guilford Press, 2001, p. 20.

Vandenberg S. G., Kuse A. R., "Mental rotations, a group test of three-dimensinal spatial visualization", *Percept Mot Skills*, Vol. 47, 1978, pp. 599-604.

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

Westerberg H., Jacobaeus H., Hirvikoski T., Clevberger P., Östensson M.-L., Bartfai A., Klingberg T. "Computerized working memory training after stroke – A pilot study", *Brain Injury*, Vol. 21, No. 1, 2007, pp. 21-29.