EVIDENCE-BASED ASSESSMENT AND PSYCHOLOGICAL INTERVENTIONS DOCTORAL SCHOOL
BABEȘ-BOLYAI UNIVERSITY

Ph.D. THESIS

THE VALIDITY OF VIRTUAL REALITY-BASED NEUROPSYCHOLOGICAL ASSESSMENT MEASURES

EXTENDED ABSTRACT

AUTHOR: Ph.D. CANDIDATE NEGUȚ ALEXANDRA
SCIENTIFIC ADVISOR: PROFESSOR Ph.D. DAVID DANIEL

CLUJ-NAPOCA
2015

TABLE OF CONTENTS

Keywords: virtual reality, ecological validity, neuropsychological assessment, test validity, task difficulty

CHAPTER 1. THEORETICAL BACKGROUND

Introduction and research topic[1]

Virtual reality technology is based on an advanced human-computer interface which generates a 3D environment and uses a wide range of technologies such as: trackers and head mounted displays (HMDs) which supply the visual input, headphones and gesture-sensing gloves for acoustic input; as well as data gloves or joysticks which provide and enhance interaction. By using these devices among with appropriate software the person is immersed into a virtual environment generated by the computer (Elkind, Rubin, Rosenthal, Skoff, & Prather, 2001; Parsons, 2012).

Recently, virtual reality scenarios emerged as a promise tool in neuropsychological assessment (Brooks & Rose, 2005; Parsons, 2012; Rizzo, Schultheis, Kerns, & Mateer, 2004; Rose, Brooks, & Rizzo, 2005; Schultheis, Himelstein, & Rizzo, 2002) and rehabilitation of cognitive processes (Foreman & Stirk, 2005; Man, 2010; Rose et al, 2005) and in clinical psychology as part of the desensitization process used in the treatment of different phobias such as: acrophobia, agoraphobia, claustrophobia, fear of flying and fear of public speaking (Bullinger, Roessler, & Mueller-Spahn, 2000; Kahan, Tanzer, Darvin, & Borer, 2000; North, North, & Coble, 1995; Mühlberger et al., 2001; Mühlberger, Sperber, Wieser, & Pauli, 2008; Mühlberger, Weik, Pauli, & Wiedemann, 2006; Rothbaum, Hodges, Kooper, Opdykes, Williford, & North, 1995; Vincelli, Choi, Molinari, Weiderhold, & Riva, 2000). Furthermore,

---

[1] This study was published.

Neguţ, A. (2014). Cognitive assessment and rehabilitation in virtual reality: theoretical review and practical implications. *Romanian Journal of Applied Psychology,* 16(1), 1-7.

virtual reality applications is expanding to clinical uses in driving assessment for persons with brain injury (Schultheis, & Mourant, 2001; Wald, Liu, & Reil, 2000), in training people with learning difficulties (Lannen, et al., 2002) or intellectual disabilities (Standen, & Brown, 2005).

Although virtual reality represents a relative new area of research and practice in the psychology field, advances in technology and computer science have supported the development of more accessible and usable virtual reality systems. As a consequence, the costs of virtual reality devices have been reduced. In addition, technical and software features of virtual reality environments are easily modified so that it allows multiple applications from which various target populations may benefit from (Elkind et al., 2001; Rizzo et al., 2006).

Because of its advantages, virtual reality environments are a promising tool in cognitive assessment and rehabilitation. Nevertheless, there is need for more studies carried out for different types of cognitive processes, conducted on different clinical population, and with different measurement instruments, not only to validate virtual reality measures, but also, to develop new procedures and interventions for a more reliable and ecological assessment. Therefore, in the first chapter we will present the state of the art of research conducted on the topic of virtual reality-based neuropsychological assessment. We will focus on the three main approaches on neuropsychological assessment: classical paper-and-pencil neuropsychological psychometrics, computerized neuropsychological tests, as well as virtual reality-based neuropsychological tests. Afterwards, we will discuss the current direction and research concerning the validity of virtual reality-based neuropsychological assessment applied on attention processes and memory. Further on, in Chapter II we will describe the main research objectives followed by the methodology used to achieve our objectives. In order to examine the validity of virtual reality measures in neuropsychological assessment in Chapter III we will present our original research. The original research consists of four studies: one meta-analysis, followed by three experimental studies. Chapter IV will include the summary of the results. Also, issues concerning theoretical and practical implications of our results will be addressed, as well as potential limitations and future research directions will be discussed.

Validity issues in psychological testing

When addressing the validity of a neuropsychological measurement instrument, in particular, and in psychological testing, in general, one crucial issue concerns validity issues. Current approaches on psychological testing consider that validity is not a characteristic of a test per se. Instead a more preferred direction discusses validity in light of the validity of test scores. It is considered that the validity of test scores describe the amount of evidence that can be brought to support the inferences made on the basis of test results (Urbina, 2004). Further on, the literature on psychological testing (Urbina, 2004) proposes several sources of validity evidence and validation studies should focus on gaining evidence in favor of test validity from as many sources as possible in relationship to the purpose of the test. Due to the fact that construct validity is referred as the core aspect of test scoring and interpretation we focused on this type of validity. Even more, most forms of validity, such as content and criterion-related, as well as patterns of convergence and divergence are considered as part of construct validity (Urbina, 2004; Messick, 1995).

Ecological validity

The topic of ecological validity has become an important topic of discussion in neuropsychological assessment. Ecological validity implies a close link between the challenges imposed by the assessment procedures and the challenges that the subject has to confront in real life situations (Wasserman & Bracken, 2003). In a broad sense, ecological validity of a test refers to its' capacity to provide similar results with those expressed in real life situations (Chaytor & Schmitter-Edgecombe, 2003; Wasserman & Bracken, 2003). An adequate level of ecological validity of a neuropsychological test implies that on the basis of the test results and on their interpretation one can make accurate predictions of behavior in real life situations (Van der Elst et al., 2008).

Two approaches to the issue of ecological validity are addressed in the field: verisimilitude and veridicality. Verisimilitude focuses on the degree to which the demands and challenges of the neuropsychological test resemble to those expressed in real life settings (Chaytor & Schmitter-Edgecombe, 2003). For instance an attempt to increase the verisimilitude of a test requests the inclusion of everyday cognitive tasks in the test items. The main focus of such approach is not on the discriminant utility of a test, but on how well the test scores reflect every day cognitive skills. The purpose of tests with increased verisimilitude is not to detect and discriminate people with brain damage, but to identify the persons that are impaired on everyday activities (Chaytor & Schmitter-Edgecombe, 2003). In the attempt to create tests with verisimilitude several measures have been developed. For executive function assessment the Behavioral Assessment of the Dysexecutive Syndrome (Wilson et al., 1986) has been developed. In attention processes assessment the Test of Everyday Attention (Robertson et al., 1996) is used, and the Rivermead Behavioral Memory Test for memory (Wilson et al., 1985). In line with the verisimilitude approach these tests simulate daily cognitive tasks: to remember the location of predefined objects or to plan a route.

Veridicality quantifies the magnitude of the relationship between test results and objective criteria of real world performance. This approach to ecological validity of neuropsychological tests relies more on statistical procedures that tests the degree to which scores from the tests are related to real life criteria (e.g. academic performance, employment status) (Chaytor & Schmitter-Edgecombe, 2003; Spooner, 2006).

Overall, results from tests that examined the ecological validity of classical psychometrics in clinical settings (Chaytor & Schmitter-Edgecombe, 2003; Chaytor, Schmitter-Edgecombe, & Burr, 2006), as well as on healthy populations (Spooner, & Pachana, 2006; Van der Elst, Van Boxtel, Van Breukelen, & Jolles, 2008) point out a low to moderate level of ecological validity in predicting real world functioning. There is an increasing need for neuropsychological assessment to assess what a subject does in everyday activities and not what he can do (Chaytor & Schmitter-Edgecombe, 2003).

Relevance and impact of the current research

Dysfunctions in the central nervous system might cause cognitive and functional impairments. Usually, these impairments reflect in cognitive processes, such as: attention, memory, language, spatial abilities, higher reasoning, functional abilities and executive function (Elkind, Rubin, Rosenthal, Skoff, & Prather, 2001; Rizzo et al., 2000). Conditions responsible for CNS dysfunction are: traumatic brain injury, stroke, Alzheimer's disease, vascular dementia, Parkinson's disease, Huntington's disease, cerebral palsy, epilepsy and multiple sclerosis (Rizzo et al., 2000).

Current tools used in the assessment of cognitive functioning rely on classical paper-and-pencil psychometrics or computer-delivered continuous performance tests and consist of certain amount of stimuli delivered to the subjects in a highly systematic and controlled environment. The gold standard in neuropsychological assessment is classical paper-and-pencil tests and computerized tests. These tests have good psychometric properties and construct validity. However, research has pointed out that classical paper-and-pencil tests have a low to moderate level of ecological validity (Chaytor & Schmitter-Edgecombe, 2003; Chaytor, Schmitter-Edgecombe, & Burr, 2006; Spooner, & Pachana, 2006; Van der Elst, Van Boxtel, Van Breukelen, & Jolles, 2008). Therefore, it is recommended to develop other assessment instruments with larger levels of ecological validity (Alvarez & Emory, 2006; Chaytor & Schmitter-Edgecombe, 2003; Elkind et al., 2001; Schultheis et al., 2002). In other terms, neuropsychological assessment should move one step further from measuring what a person can do and to replace it with an assessment that describes what the person actually does in the real world (Chaytor & Schmitter-Edgecombe, 2003). Virtual reality-based neuropsychological assessment might be considered as an alternative to classical paper-and-pencil tests or computerized measures with an increased level of ecological validity. Virtual reality-based neuropsychological assessment is the systematic procedure that contains cognitive tasks embedded in a virtual reality scenario for the evaluation of specific activities in the central nervous system (CNS) that are associated with observable behaviors.

Because of its unique features that create the premises of an ecological neuropsychological evaluation virtual reality environments represent promising tools in neuropsychological assessment there is need for further investigation. First of all, in order to achieve support for construct validity, convergent and divergent validity and predictive validity and to empirically validate the constructs more studies based on virtual reality environments have to be carried out by different research teams. Second, there is a lack of investigations on the mechanisms by which virtual reality works. Research should focus on underlying the way in which human factors interact with virtual reality technology.

CHAPTER II. RESEARCH OBJECTIVES AND OVERALL METHODOLOGY

As pointed out in the first chapter that highlights the need to develop neuropsychological measurement instruments with an increased ecological validity because current assessment tools have a low to moderate level of ecological validity we aim to investigate the evidence that can be brought in favor of the validity of a relative new approach to neuropsychological assessment. This new approach is virtual reality-based neuropsychological assessment. Further on, in the current chapter we will present our research objectives, as well as the subsequent methodology that was applied to achieve our objectives.

The general aim of this Ph.D. thesis is to investigate the validity of assessment using virtual reality by running several studies that might be considered evidence for different forms of validity. In order to discuss and investigate the validity of virtual reality-based neuropsychological assessment in line with recommendation from the literature, we concentrated on content-related, criterion-related validity, and patterns of convergence. One source of content-related validity that was targeted in the current research results by investigating the association between distinct brain activities expressed in observable behaviors and the neurological foundations that underlie those activities. Content-related validity can be achieved by linking the relevance and representativeness of test content with test responses. Next, criterion-related forms of validity were addressed by investigating the diagnostic validity of the virtual reality-based neuropsychological assessment tool. This type of validity is established by comparing two contrasted groups, coming from clinical and healthy populations. Superior performance of the healthy control group is considered as evidence in favor of the diagnostic validity of the test. Large positive correlations between two tests measuring the same theoretical constructs demonstrate convergent validity (Urbina, 2004; Wasserman & Bracken, 2003). Addressing these issues in regard to validity might be considered an important theoretical advance when investigating and discussing the validity of a test. Because virtual reality-based neuropsychological assessment can be considered as a relative new topic of research in applied neuropsychology there is need to argue on its validity.

CHAPTER III. ORIGINAL RESEARCH

Study 1. Validity of virtual reality neurocognitive assessment: a meta-analytic approach[2]

Introduction

Neuropsychological assessment is considered an applied science that focuses on the evaluation of specific activities in the central nervous system (CNS) that are associated with observable behaviors (Lezak, 1995). Neuropsychological evaluation is performed with different types of measurement instruments which are standardized and have good psychometric properties, such as reliability and validity (Morganti, 2004; Schultheis et al., 2002). Central

[2] Parts of this study were published, accepted for publication or submitted for publication.

Neguţ, A., Matu, S. A., Sava, F. A., & David, D. (2015). Convergent validity of virtual reality neurocognitive assessment: a meta-analytic approach. *Erdelyi Pszichologiai Szemle= Transylvanian Journal of Psychology*, *16*(1), 31.

Neguţ, A., Matu, S. A., Sava, F. A., & David, D. (in press). Task difficulty of virtual reality-based assessment tools compared to classical paper-and-pencil or computerized measures: A meta-analytic approach. *Computers in Human Behavior*, *54*, 414-424. doi: 10.1016/j.chb.2015.08.029

Neguţ, A., Matu, S. A., Sava, F. A., & David, D. (submitted). Virtual reality measures in neuropsychological assessment: a meta-analytic review. *The Clinical Neuropsychologist* nervous system dysfunctions result in cognitive and functional impairments that imply processes of attention and executive functions, memory and learning, language, spatial abilities, higher reasoning, functional abilities (Elkind et al., 2001; Rizzo et al., 2000; Schultheis et al., 2002). Various conditions are responsible for CNS dysfunction, such as: traumatic brain injury, stroke, Alzheimer's disease, vascular dementia, Parkinson's disease, Huntington's disease, cerebral palsy, epilepsy and multiple sclerosis (Rizzo et al., 2000).

Classic paper-and-pencil psychometrics (Brandt & Benedict, 2001; Benedict, 1997; Delis, Kramer, & Kaplan, 2001; Delis, Kramer, Kaplan, & Ober, 1987; Halligan, Marshall, & Wade, 1989; Heaton, Chelune, Talley, Kay, & Curtiss, 1993) as well as computer-based assessment instruments (Conners, 2000; Reeves, Kane, Winter, & Goldstone, 1995; Greenberg & Waldman, 1993; Verbruggen, Logan, & Stevens, 2008) represent the current standard assessment tools used in neuropsychological evaluation. They consist of a certain amount of stimuli delivered to the subjects in a highly systematic and controlled environment via written paper or a computer screen. Also, scoring and test interpretation are conducted either by a trained practitioner or automatically by the computer (Bauer et al., 2012; Butcher, 2003; Kane & Kay, 1992; Podell, DeFina, Barrett, McCullen, & Goldberg, 2003).

A new paradigm for neuropsychological assessment is virtual reality-based assessment. Virtual reality instruments can now be used for the cognitive assessment of executive functions, attention, and impulsivity, cognitive and motor inhibition (Elkind et al., 2001; Henry et al., 2012; Henry et al., 2011; Ku et al., 2003; Parsons, Courtney, Arizmendi, & Dawson, 2011; Pugnetti, Mendozzi, Barbieri, & Motta, 1998b; Rizzo et al., 2000), memory and learning (Gamberini, 2000; Matheis, Schultheis, Tiersky, DeLuca, Millis, & Rizzo, 2007; Parsons & Rizzo, 2008b;

Pugnetti et al., 1998a), spatial abilities (Parsons et al., 2004) and visuospatial neglect (Broeren, Samuelsson, Stibrant-Sunnerhagen, Blomstrand, & Rydmark, 2007).

Overview of the current study

Despite the fact that previous research has provided a useful database on the topic of virtual-reality based neuropsychological assessment, no meta-analysis has been conducted in order to investigate the validity of virtual reality measures. Validity can be defined as "what the test measures and how well it does so" (Anastasi & Urbina, 1997, p. 113) and "hinges on the evidence we can bring to bear to support any inference that is to be made on the basis of test results" (Urbina, 2004, p. 151). Validity is the core aspect of test scoring and its use. Although a reasonable number of theoretical reviews have been published on this topic (Elkind, 1998; Myers & Bierig, 2000; Rizzo et al., 1999; Riva, 1998), there is no systematic review on neuropsychological assessment using virtual reality. Thus, we conducted a meta-analysis to examine the validity of neuropsychological assessment in virtual reality in comparison with traditional paper-and-pencil and computerized assessments. Nevertheless, giving that virtual reality neuropsychological assessment techniques are spreading in both scientific and clinical communities, and their potential benefits over classical and computerized measures, a meta-analysis could help clarify important issues regarding their discriminant validity, task difficulty, and complexity.

The current meta-analysis sought to examine the following objectives:

1. To investigate differences in performance on virtual reality-based measures of cognitive processes between clinical and healthy populations;

2. To investigate the strength and the direction of the relationship between classical paper-and-pencil or computerized measures and virtual reality-based measures;

3. To examine differences in performance between classical or computerized measures and virtual reality-based measures of cognitive processes;

4. To investigate potential moderators of the results.

In our approach, we focused on construct validity because it is a core aspect of test scoring and interpretation. According to current directions on psychological testing, construct validity is a fundamental issue regarding test validity because most forms of validity, such as content and criterion-related, as well as patterns of convergence and divergence are considered as part of construct validity (Urbina, 2004; Messick, 1995).

Method

Literature search

In order to identify potentially relevant studies, a systematic literature search on virtual reality assessment has been conducted using “virtual reality” and “cognitive assessment” as search terms in Medline, PsychInfo and ScienceDirect databases, up to June 2014. Furthermore, the list of references of empirical articles and reviews on this topic were screened in order to detect other studies that did not appear in the electronic search.

Studies selection

The following criteria were used for the inclusion of studies in the meta-analysis: (a) any experimental study with minimum two conditions: one using a virtual reality-based measurement and one using a classical or computerized measurement of the same cognitive process; (b) any experimental study with minimum two experimental groups: a control group and a clinical group measured with the same virtual reality assessment tool; (c) any study which included correlational analysis between classical or computerized measurement instruments and virtual reality assessment tools of the same cognitive process; (d) assessed any neuropsychological process using virtual reality and analogous classical or computerized assessment tools; (e) provided sufficient data to compute effect sizes; (f) were English-based publications.

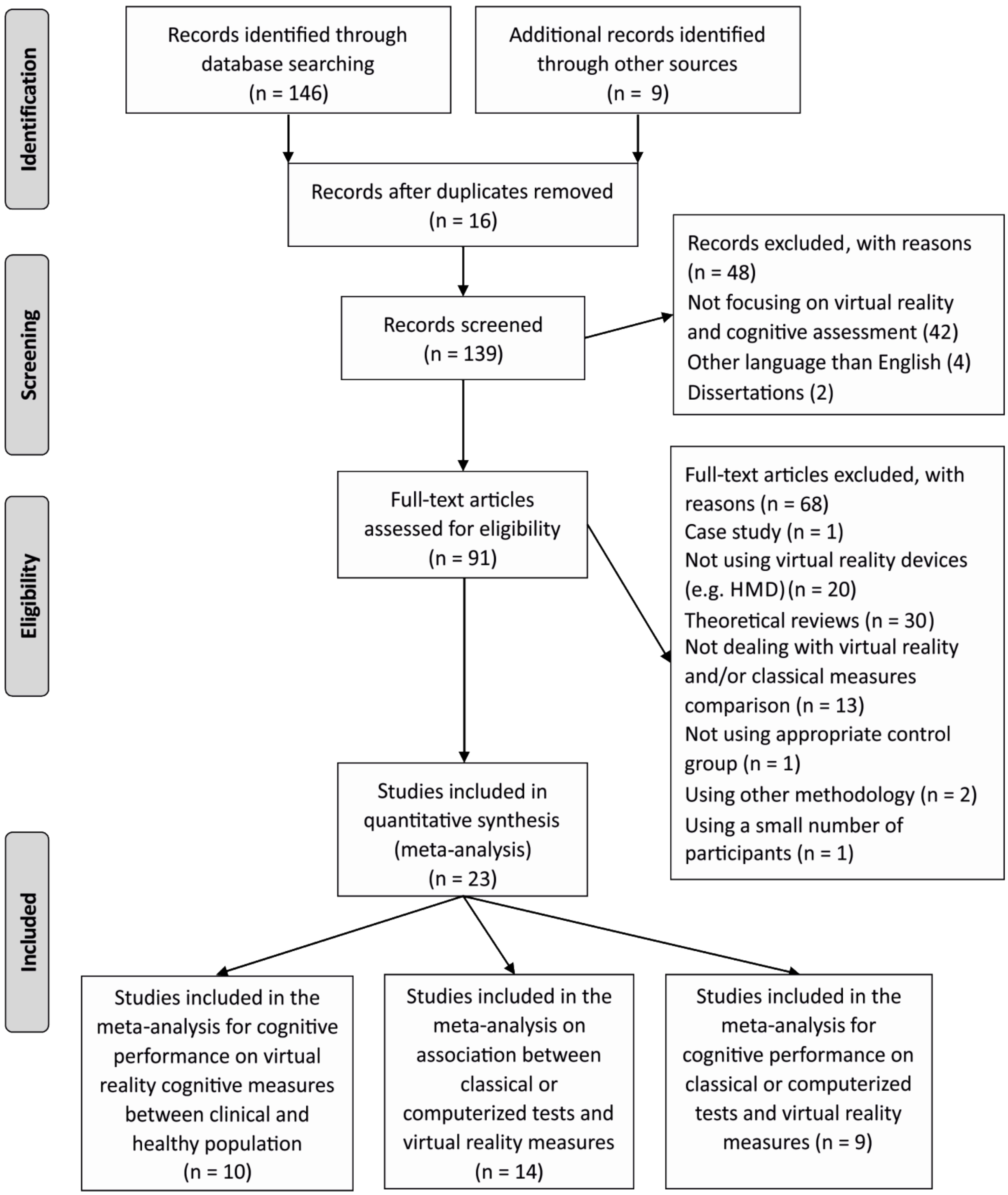

Figure 1. PRISMA flow diagram

Data coding

Outcome measures were classified into three categories based on the cognitive process assessed, and subsequent cognitive assessment scales: executive functions, memory, and other neurocognitive measures. Only these measures were available for analysis from the studies that met the inclusion criteria in the meta-analysis.

Effect size calculation and heterogeneity

For the first objective (differences in performance on virtual reality-based measures of cognitive processes between clinical and healthy populations) and third objective (differences in performance between classical or computerized measures and virtual reality-based measures of cognitive processes), between-group effect sizes were calculated using Hedges's *g*. To address the second objective (the strength and the direction of the relationship between classical paper-and-pencil or computerized measures and virtual reality-based measures), we used the correlation coefficient *r* as an effect size index using Borenstein et al. (2009) approach.

Results

Comparison of the performance between clinical and control groups on virtual reality-based measures

For the first objective, we computed average effect sizes from 10 studies comparing performance of the clinical and control groups on virtual reality measures (N = 348). We found large mean effect in favor of the healthy control group ($g = 1.21$, 95% CI [0.77, 1.65], $z = 5.39$; $p < .001$). Nevertheless, there was evidence of heterogeneity in the results ($Q_{(9)} = 30.76$, $p < .001$; $I^2 = 70.74\%$) which was addressed by performing moderation analysis.

Meta-analysis for performance on virtual reality-based cognitive measures between clinical and healthy population

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Adams et al., 2009	0.592	0.346	0.120	-0.087	1.270	1.709	0.088
Broeren et al., 2007	1.651	0.688	0.473	0.303	2.999	2.401	0.016
Kang et al., 2008	1.485	0.352	0.124	0.795	2.174	4.221	0.000
Kim et al., 2004	1.718	0.373	0.139	0.987	2.449	4.605	0.000
Ku et al., 2004	1.003	0.363	0.132	0.292	1.714	2.765	0.006
Ku et al., 2003	1.000	0.404	0.164	0.207	1.792	2.472	0.013
Matheis et al., 2007	2.954	0.462	0.214	2.049	3.860	6.393	0.000
Parsons et al., 2007	1.200	0.472	0.223	0.275	2.125	2.541	0.011
Pugnetti et al., 1998	0.467	0.256	0.066	-0.035	0.969	1.823	0.068
Siemerkus et al., 2012	0.644	0.355	0.126	-0.051	1.339	1.815	0.069
	1.214	0.225	0.051	0.773	1.655	5.398	0.000

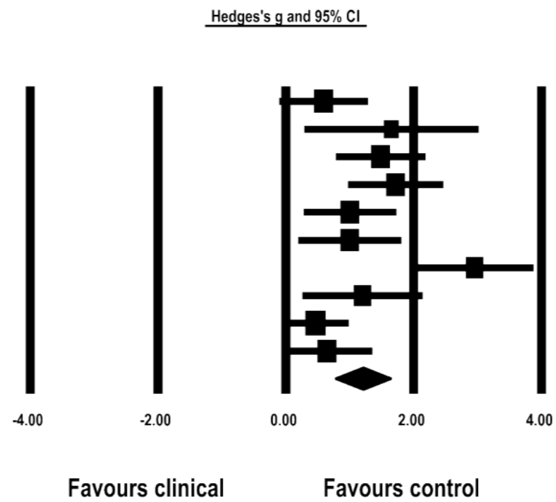

Figure 3. Meta-analysis for performance on virtual-reality-based cognitive measures between clinical and healthy population

For the second objective, we computed mean effect sizes from fourteen studies (N= 553) using correlation coefficient *r*. Overall, results point out a positive significant medium correlation between virtual reality measures and classical or computerized measures(r = .51, 95% CI [0.27, 0.68], z = 3.92; p < .001) with high significant heterogeneity ($Q_{(13)}$ = 123.83, p < .001; I^2 = 89.50%). The value of the effect size indicates a moderate to good relationship.

Meta-analysis for association between classical or computer-based measures and virtual reality-based measures

Study name	Correlation	Lower limit	Upper limit	Z-Value	p-Value
Adams et al., 2009	0.378	-0.111	0.719	1.532	0.126
Armstrong et al., 2012	0.492	0.245	0.679	3.653	0.000
Elkind et al., 2001	0.404	0.173	0.592	3.318	0.001
Henry et al., 2012	0.156	-0.080	0.376	1.300	0.194
Ku et al., 2003	-0.340	-0.643	0.054	-1.699	0.089
Lalonde et al., 2013	0.324	0.005	0.583	1.990	0.047
Matheis et al., 2007	0.700	0.497	0.830	5.276	0.000
Parsons & Rizzo., 2008	0.571	0.265	0.772	3.373	0.001
Parsons et al., 2004	0.502	-0.145	0.849	1.551	0.121
Parsons et al., 2007a	0.498	0.071	0.771	2.254	0.024
Parsons et al., 2007b	0.569	0.168	0.808	2.661	0.008
Parsons et al., 2008	0.675	0.460	0.815	4.987	0.000
Parsons et al., 2013	0.352	0.082	0.574	2.524	0.012
Parsons et al., 2014	0.955	0.922	0.974	12.925	0.000
	0.511	0.275	0.689	3.924	0.000

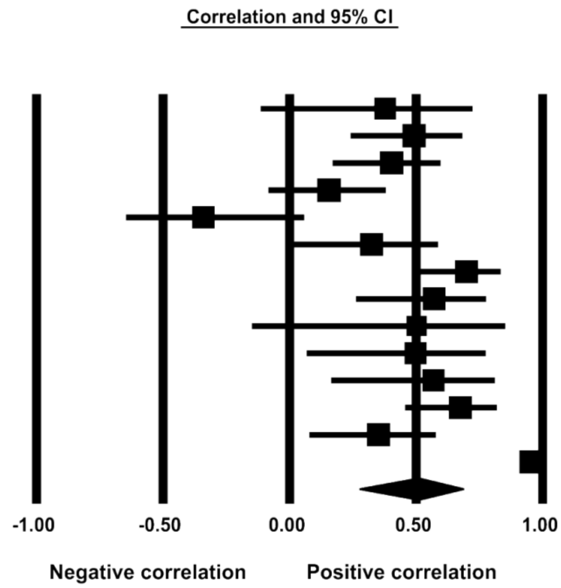

Figure 4. Meta-analysis for association between classical or computer-based measures and virtual reality-based measures

For the third objective the average effect sizes were calculated from nine studies (N = 301), one that used a between-subject design (Gamberini, 2000), and eight that used a within-subject design. The resulted effect size was adjusted using Olejnik & Algina's (2000) technical specifications. Results showed significant differences between virtual reality measures and computerized or classical measures with a medium effect size in favor of classical or computerized measures (g = -0.77, 95% CI [-1.50, -0.05], z = -2.09; p = .036). There was also evidence of high heterogeneity ($Q_{(8)} = 125.56$, $p < .001$; $I^2 = 93.62\%$). The negative sign indicates that classical or computerized assessments yield better performance.

Meta-analysis for cognitive performance on classical or computer-based measures and virtual reality-based measures

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Armstrong et al., 2012	-2.529	0.326	0.106	-3.167	-1.890	-7.765	0.000
Broeren et al., 2007	-0.184	0.369	0.136	-0.907	0.540	-0.497	0.619
Elkind et al., 2001	-0.515	0.134	0.018	-0.778	-0.253	-3.843	0.000
Gamberini, 2000	-0.459	1.109	1.231	-2.633	1.716	-0.413	0.679
Nolin et al., 2009	-1.170	0.496	0.246	-2.142	-0.198	-2.360	0.018
Parsons et al., 2013	-4.652	0.944	0.890	-6.501	-2.802	-4.930	0.000
Parsons et al., 2014	-0.445	0.149	0.022	-0.738	-0.152	-2.975	0.003
Pollak et al., 2010	-0.645	0.218	0.048	-1.072	-0.217	-2.958	0.003
Pugnetti et al., 1998	2.091	0.323	0.104	1.457	2.724	6.468	0.000
	-0.776	0.370	0.137	-1.501	-0.050	-2.096	0.036

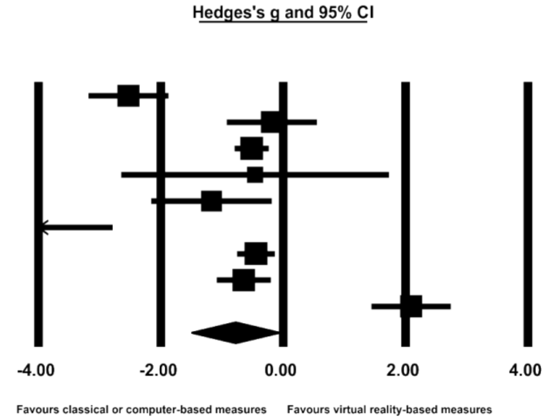

Figure 4. Meta-analysis for cognitive performance on classical or computer-based measures and virtual-reality-based measures

Moderation analysis

To investigate the fourth objective, we conducted moderation analysis for each of the three types of analysis: between-group analysis (clinical versus healthy population) of the performance on virtual reality-based cognitive measures, correlational analysis between virtual reality measures and classical or computerized tests, and between-group analysis for cognitive performance on virtual reality measures and classical or computerized tests.

Table 6

Moderation analysis with categorical variables for performance on virtual reality cognitive measures

Outcome	Moderator	*K*	*g*	*p*	Q w	*p*	*95%* CI	Q b	*p*
Performance on cognitive measures	Schizophrenia/	3	0.78	.000	3.41	.906	[0.54; 1.02]	4.40	.110
	brain injury/	5	1.09	.000	124.86	.000	[0.93; 1.25]		
	ADHD	2	0.99	.000	50.51	.011	[0.84; 1.15]		
	Active exploration/	6	0.77	.000	35.74	.003	[0.61; 0.93]	12.97	.000
	Passive exploration	4	1.15	.000	134.48	.000	[1.02; 1.28]		
	Time-based measures/	3	0.98	.000	26.60	.000	[0.70; 1.27]	0.33	.562
	Error-based measures	10	0.89	.000	137.73	.000	[0.77; 1.02]		
	Distractors/	4	0.93	.000	43.43	.013	[0.77; 1.09]	0.99	.318
	No distractors	7	1.03	.000	138.77	.000	[0.90; 1.16]		

Note. *K* = number of studies included in the analysis; *g* = Hedge's *g*; 95% CI = 95% confidence interval around the weighted mean effect size.

Table 8

Moderation analysis with categorical variables for association between classical or computerized tests and virtual reality measures

Outcome	Moderator	*K*	*r*	*p*	Q w	*p*	*95%* CI	Q b	*p*
Correlation coefficient between classical or computerized tests and virtual reality measures	Classic/	13	.30	.000	612.04	.000	[0.27; 0.33]	11.59	.001
	Computer	4	.22	.000	164.08	.000	[0.17; 0.26]		

Note. *K* = number of studies included in the analysis; *r* = correlational coefficient; 95% CI = 95% confidence interval around the weighted mean effect size.

Table 10

Moderation analysis with categorical variables for performance on cognitive measures

Outcome	Moderator	*K*	*g*	*p*	Q w	*p*	*95%* CI	Q b	*p*
Performance on cognitive measures	Healthy/	6	-0.71	.000	421.02	.000	[-0.79; -0.63]	5.74	.017
	Clinic	3	-0.52	.000	40.18	.000	[-0.65;-0.38]		
				.000					
	Classic/	6	-0.58	.000	269.59	.000	[-0.66; -0.50]	10.15	.001
	Computer	5	-0.82	.000	187.20	.000	[-0.95; -0.70]		
	Time-based measures/	5	-1.07	.000	269.54	.000	[-1.19; -0.95]	64.59	.000
	Error-based measures	8	-0.46	.000	132.81	.000	[-0.54; -0.38]		

Note. *K* = number of studies included in the analysis; *g* = Hedge's *g*; 95% CI = 95% confidence interval around the weighted mean effect size.

Table 5

Mean effect sizes (Hedges's g) for overall cognitive performance, executive functions, memory and other neurocognitive measures depending on type of samples included in the analysis for the comparison between virtual reality measures and classical or computerized measures

Outcome	Effect size for clinical population	Effect size for healthy population
Overall cognitive performance	-0.71 (*K* = 3)	-0.85 (*K* = 6)
Executive functions	-0.79 (*K* = 2)	-1.07 (*K* = 4)
Memory	(*K* = 0)	1.00 (*K* = 2)
Other neurocognitive measures	-0.21 (*K* = 1)	(*K* = 0)

Note. K = number of studies included in the analysis; Effect size = Hedges's *g*

Publication bias

We used Duval and Tweedie's trim-and-fill procedure in order to investigate the presence of publication bias for each of the three main analyses corresponding to our objectives. For the comparison of clinical and healthy populations on virtual reality based measures, trim-and-fill procedure identified one study with an effect higher than the mean which increased the effect size value, but it did not modify its' magnitude (*g* = 1.32, 95% CI = [0.86; 1.78], Q = 45.46.

In case of the second objective, trim-and-fill procedure estimated five studies with an effect size higher than the mean which did not modify the magnitude of the effect size (after including the five hypothetical studies *r* = .65, 95% CI = [0.47; 0.78], Q = 216.82).

For the third objective, trim-and-fill procedure estimated that no studies are missing which could modify the results.

Discussion

The present meta-analysis investigated the validity of virtual reality-based measures in comparison to classical paper-and-pencil or computerized neuropsychological measures. In line with current approaches on validity issues, we aimed to provide evidence that support the use of virtual reality assessment procedures. It is noteworthy that the evidence of test validity may take many different forms (Urbina, 2004; Wasserman & Bracken, 2003). The present research dealt mainly with construct validity. Each type of analysis focused either on diagnostic validity, convergent validity and task difficulty of virtual reality-based measures as relevant evidence for test sensitivity and specificity.

Overall, results from the current meta-analysis bring evidence in favor of the validity of virtual reality-based measures. It is important to notice that virtual reality-based tests have discriminant validity as they discriminate between healthy and control participants. In contrast to classical or computerized measures, virtual reality-based instruments have an increased level of task difficulty as they are more demanding and require additional cognitive resources. Based on this argument we speculate that virtual reality based measures appear to be more appropriate for mild cognitive impairments. Test scores from virtual reality assessment correlate moderately with current neuropsychological assessment tools, which suggest that the two types of assessment tools tap to some extent into their results the same cognitive processes. To conclude, virtual reality assessment tools appear to be valid instruments for cognitive assessment.

Study 2. Virtual reality-based attention assessment in comparison with computerized assessment in ADHD: the Virtual Classroom versus an analogue Continous Performance Test[3]

Introduction

For ADHD diagnosis, common measures include parent or teacher behavioral rating scales as well as clinical interviews. Although they are critical to a valid assessment, they have limited predictive validity. In plus, their treatment utility is reduced, because they do not target specific cognitive mechanisms that underlie the attention deficit (Losier, McGrath, & Klein, 1996; Parsons, Bowerly, Buckwalter, & Rizzo, 2007). Even more, the use of rating scales alone may lead to overdiagnosis (Weiler et al., 2000). To overcome these issues and to increase the

[3] Neguţ, A., Jurma, A. M., & David, D. (submitted). Virtual reality-based attention assessment in comparison with computerized assessment in ADHD: the Virtual Classroom versus an analogue Continous Performance Test. *Child Neuropsychology*

cost effectiveness of neuropsychological assessment of ADHD it is recommended the use laboratory-based measures or neuropsychological tests (Losier et al., 1996; Parsons et al., 2007).

One of the most valid, reliable and used measure of sustained vigilance, attention and impulsivity is the Continuous Performance Test (CPT). The outcomes of interest for attention assessment with CPTs are total correct responses, errors of commission, errors of omission, number of hits recorded, and the mean reaction time (Barkley, Grodzinsky, & DuPaul, 1992; Bioulac et al., 2012; Frazier et al., 2004; Gilboa et al., 2011; Losier et al., 1996; Parsons et al., 2007). Overall, results from meta-analysis comparing performance between children with ADHD and healthy controls point out medium to large effect sizes for omission errors, commission errors and standard deviation of reaction time and medium effect sizes for reaction time (Huang-Pollock, Karalunas, Tam, & Moore, 2012; Losier et al., 1996).

Virtual Classroom (Rizzo et al., 2006; Rizzo et al., 2000) is a neuropsychological test embedded in virtual reality designed to assess attention deficits in children with ADHD or other conditions associated with impaired attention like neurofibromatosis type 1 (Gilboa et al., 2011) or traumatic brain injury (Nolin, Martin, & Bouchard, 2009). Few studies have investigated the diagnostic validity of Virtual Classroom by comparing attention performance of children with ADHD and healthy controls.

Overview of the current study

Because results from the studies that investigated the discriminant validity of Virtual Classroom in comparison to CPTs point out mixed results, the aims of the current study are to (1) investigate the discriminant validity of Virtual Classroom in attention assessment by comparing performance of children with ADHD with healthy controls, (2) explore the task difficulty of virtual reality-based measures by comparing attention performance obtained with a virtual reality-based measure and an analogue CPT, (3) to address the effect of distractors on performance of ADHD participants and healthy controls, and (4) to compare the two measures on cognitive absorption.

Method

Participants

Forty-five boys and thirty girls aged between 7 and 13 years old (m = 9.49, SD = 1.67) participated in the study. There were differences on age between the two groups $t(55)=3.56$, $p < .001$ (mean age of the control group = 8.9 vs. mean age of ADHD group = 10.24).

Measures

Sociodemographic variables. Parents or caregivers reported children age, gender, psychiatric diagnostic and pharmacological treatment, and eye problems.

Several cognitive measures that seem to discriminate between children with ADHD and healthy children were chosen (Frazier et al., 2004).

Executive function measures. The Digit Span and Letter Number Sequencing subtests from the WISC-IV (Wechsler, 2003).

General intelligence. The Romanian form of Raven Standard Progressive Matrices Plus (Dobrean, Raven, Comşa, Rusu, & Balázsi, 2008; Domuța, Balázsi, Porumb, Rusu, & Comşa, 2003) was used in this study to measure intelligence.

The Simulator Sickness Questionnaire (SSQ, Kennedy, Lane, Berbaum, & Lilienthal, 1993) was administrated to children in order to determine any sickness symptoms due to immersion in Virtual Reality.

An adapted version for children of Cognitive Absorption Scale (CAS, Agarwal & Karahanna, 2000) was used to measure the state of deep involvement with software which predicts usage behavior.

Virtual Classroom is a virtual reality-based attention processes measure developed by Rizzo et al. (2006; 2000) with Digital MediaWorks Inc. (http://web.dmw.ca/). It follows a classical continuous performance test scenario in which the participant is exposed to stimuli over a long period of time and has to respond as quickly as possible to target stimuli and to inhibit any responses to non-target stimuli. The Virtual Classroom scenario consists of a rectangular classroom populated by desks, a blackboard, windows and doorways on each side of the classroom, pupils and a female teacher in front of the classroom.

The CPT used in this research replicated the stimulus challenges from the Virtual Classroom without immersion into the classroom. To be more specific, the number of targets and non-targets, the total number of targets, as well as the inter-stimulus interval and duration were identical with the Virtual Classroom scenario. The CPT was designed using Inquisit 3 Software (2012).

The following measures from the Virtual Classroom and analogue CPT were used: total correct responses, errors of commission, errors of omission, and the mean reaction time.

Procedure

Written consent from the parents or legal guardians was obtained before testing. Children gave an oral consent and were informed that they can stop the testing if they need to without penalty. The same neuropsychological tests were administered to all children, with the difference that half of them received the Virtual Classroom assessment while the other half were tested using the CPT. Prior to testing all of the participants received an ID and were randomly assigned to one of the experimental conditions (assessment using Virtual Classroom or CPT). Each of the participants was tested with and without distractors with either Virtual Classroom or CPT and the order of administration was counterbalanced within-subject.

The testing session lasted for approximately two hours. All the primary dependent variables represented by attention performance measured either by the Virtual Classroom or by an analogue CPT, such as: total correct responses, errors of commission, errors of omission, and the mean reaction time were recorded automatically by the computer, whereas secondary dependent variables resulted from classical paper-and-pencil assessment were obtained afterwards.

Results

To examine differences across the four dependent variables we performed a mixed Multivariate Analysis of Covariance (MANCOVA) with (1) the type of group (ADHD and healthy controls), (2) test condition (Virtual Classroom or CPT) as between factors, and (3) the test modality condition (with and without distractors) as repeated measures factors controlling for age and IQ which were set as covariates.

First of all, results indicate that age yields a significant effect over the overall attention performance, $V = 0.13$, $F(4, 66) = 2.65$, $p < .05$ while IQ does not, $V = 0.37$, $F(4, 66) = 0.63$, $p > .05$. Results from the mixed MANCOVA using Pillai's trace point out a significant main effect of clinical status on the overall performance on the number of commission, omission errors, total correct hits and reaction time, $V = 0.30$, $F(4, 66) = 7.06$, $p < .001$. Next, Sidak corrected post hoc tests showed that on Virtual Classroom ADHD children perform worse than controls on commission errors ($p < .05$), omission errors ($p < .01$), total correct responses ($p < .01$), and slower on reaction time ($p < .01$). In case of CPT children with ADHD perform worse than controls on commissions ($p < .01$) and omissions ($p < .05$), and better on total correct responses ($p < .01$). However, no significant differences between the two groups were on reaction time ($p > .05$). There was also a significant main effect of test condition on the dependent variables, $V = 0.53$, $F(4, 66) = 19.14$, $p < .001$. However, Sidak corrected post hoc tests revealed that for children with ADHD differences on commission errors, omission errors and total correct responses between assessment using Virtual Classroom and CPT are not significant ($p > .05$), but reaction time to targets was slower ($p < .01$) in the Virtual Classroom type assessment. Next, results revealed a significant main effect of test modality on the overall performance on number of commission, omission errors, total correct hits and reaction time, $V = 0.17$, $F(4, 66) = 3.44$, $p < .05$. Such a result reflects the fact that the presence or absence of distractors yields an influence over the performance obtained by both children with ADHD and healthy controls. However, Sidak post hoc tests revealed that the only significant differences between the condition with and without distractors emerged on omissions and total correct responses in Virtual Classroom ($p < .05$.).

In order to examine our fourth objective we performed independent samples *t*-test. Results point out no significant differences between the Virtual Classroom and the CPT on none of the cognitive absorption dimensions: temporal dissociation ($p > .05$), focused immersion ($p > .05$), heightened enjoyment ($p > .05$), curiosity ($p > .05$), personal innovativeness ($p > .05$), as well as perceived ease of use ($p > .05$), usefulness ($p > .05$), and behavioral intention to use ($p > .05$).

Discussion

The current study aimed to investigate the discriminant validity of a virtual reality-based measure for attention assessment in ADHD children compared to healthy controls. Also, we aimed to examine the task difficulty of Virtual Classroom compared to a well-established measure of attention delivered via computer known as CPT. Both measures, the Virtual Classroom and CPT contained a scenario with and without distractors to assess the effect of distractors over performance. It is considered that the presence of distractors can enhance the ecological validity of an assessment instrument (Parsons et al., 20007) while increasing the task difficulty because the distractors can intensify the complexity of a task.

Overall, results from the current research brings evidence in favor of the diagnostic validity of Virtual Classroom because the measure has discriminant validity as it discriminates between ADHD children and healthy controls on all CPT's parameters: total correct responses, the number of commission and omission errors and on reaction time to targets. Further on, results show that Virtual Classroom has a similar task difficulty compared to the CPT because there were no significant differences on commission and omission errors, and total correct responses between the two types of assessment instruments. However, ADHD participants showed slower reaction time rates in virtual reality, while healthy controls also had a slower reaction time in virtual reality. Nevertheless, it seems that adding distractors to the virtual environment influence the attention performance obtained by ADHD children and healthy participants on both assessment tools in case of omissions errors and total correct responses.

Study 3. The effect of learning environment on explicit and implicit memory by applying a process dissociation procedure

Introduction

Virtual reality-based neuropsychological assessment has unique features that have the potential to increase the level of ecological validity of test results. In comparison to classical paper-and-pencil or computerized measures that are characterized by a highly systematic and controlled environment, virtual reality measures propose an assessment in which cognitive tasks are embedded into a virtual environment which replicates the real world environments (Elkind et al., 2001; Neguț et al., in press; Rizzo et al., 2006; Schultesis et al., 2002). Like in real world tasks and situations, virtual environments are costly of cognitively resources because the cognitive system has to process additional information and stimuli to solve a task (Adams et al., 2009; Neguț et al., in press). In contrast, in classical assessment a less realistic approach is used when testing a cognitive process because the assessment test usually consist of stimuli delivered via written paper or a computer screen in a controlled laboratory environment. In turn, classical measures might underestimate cognitive performance. Thus, virtual reality-based assessment might offer an ecological assessment with increased task difficulty that can have discriminant validity and offer a better baseline of cognitive deficits for rehabilitation procedures. While for executive functions results from the literature point out an increased task difficulty of virtual reality-based measures, for memory measures results are mixed (Mania & Chalmers, 2001; Neguț et al., in press). In plus, for memory performance in virtual reality two opposite hypotheses can be identified. One that supports the increased task difficulty (Elkind et al., 2001; Gamberini, 2000), while the other one propose that the realistic mental images and level of presence increases the memory performance (Dinh et al., 1999; Lin et al., 2002; Mania & Chalmers, 2001).

Overview of the current research

We aim to explore the task difficulty hypothesis regarding virtual reality in memory assessment by comparing memory performance in three learning environments which differ only in the level of immersion. A first environment corresponds to classical assessment where the stimuli appear on a desktop screen and the participants have to perform the learning task in a predefined period of time. This environment corresponds to a no immersion condition with a decrease level of presence and task difficulty. The second environment consists to a desktop 3D environment whit a moderate level of immersion, presence and task difficulty. It replicates a virtual apartment in which participants navigate and have to learn as many objects as possible in a predefined period of time. The third condition which is high on the level of immersion, presence and task difficulty corresponds to the same virtual environment. The participants are immersed into the virtual environment and have to perform the same task as in the other conditions. The virtual environment is delivered via a CAVE automatic virtual environment.

A Process Dissociation Procedure (Jacoby, 1991, Jacoby et al., 1993) will be applied to measure memory performance. It offers estimates of conscious recollection and unconscious or automatic influences. Explicit and implicit memory processes have been linked to neuropsychological pathology (Baddeley & Wilson, 1994; Graf & Schacter, 1985; Jacoby, 1991; Ste-Marie et al., 1996; Rybash, & Hoyer, 1996) and are currently used in cognitive rehabilitation (Baddeley & Wilson; 1994; Dou, Man, Ou, Zheng, & Tam, 2006; Lloyd, Riley, & Powell, 2009).

Therefore, the present research sought to examine the following objectives:

1. To investigate differences between performance on explicit and implicit memory in three distinct learning environments, a computerized measure, a 3D desktop environment and 3D virtual environment.
2. To explore the strength and the direction of the relationship between cognitive absorption and explicit and implicit memory performance.
3. To examine the magnitude and direction of the association between a sense of presence into the virtual environment and explicit and implicit memory performance.
4. Investigate the association between explicit and implicit memory performance.

Method

Participants and experimental design

Seventy-seven healthy participants participated in the study aged between 19 and 39 years old ($M = 23.96$, $SD = 4.07$) out of which 22 (28.60%) were males. They were mostly psychology students that received extra credit for their participation in the study (see Table 13).

For this study we used a between-subject one-factorial experimental design with three experimental conditions that correspond to three distinct learning environments: a computerized learning environment, a 3D desktop environment and 3D virtual environment. Participants were randomly assigned to one of the experimental conditions.

Measures

Sociodemographic variables. Participants age and gender were reported.

Trail Making Test Part A and B (Reitan, 1958) was used as an executive function measure.

An adapted version of Cognitive Absorption Scale (CAS, Agarwal & Karahanna, 2000) was used as a measure of cognitive absorption.

Cybersickness was assessed with the Simulator Sickness Questionnaire (SSQ, Kennedy, Lane, Berbaum, & Lilienthal, 1993).

The level of presence was measured with Witmer and Singer Presence Questionnaire (PQ, Witmer & Singer, 1998).

Explicit and implicit memory measures were estimated using Process Dissociation Procedure (Jacoby, 1991; Jacoby et al., 1993). By applying an inclusion and exclusion test estimates of recollection and automatic or unconscious influences are obtained.

Materials

Virtual and computerized environment

The virtual environment used in the 3D desktop condition and the full immersive 3D condition was the same. The single difference consisted of the manipulation of level of immersion. The first environment was delivered on a desktop, while the other one was delivered on the Icube. The virtual scenario consisted of a virtual apartment whit 3 rooms: a bedroom, an open space living room with a kitchen and a bathroom. It contained a total of 40 objects that are typical to an apartment (e.g. bed, chair, table, glasses, bookshelf, vase, and flower) (see Appendix 2). The computerized environment contained images of the same objects within the virtual environment. The images were randomly displayed on the computer desktop screen.

Procedure

Participants were random assigned to one of the three experimental conditions in the study phase. The experiment had two main phases: study and test phase. The psychological scales related to technology and virtual reality exposer were administrated at the end of the experiment after completing the experimental procedure. In the study phase, the participants random assigned into the 3D desktop or 3D virtual reality condition were told that they will be guided through a virtual apartment. In the virtual apartment they will see different objects that are usually found into an apartment. As they will navigate into the apartment they will have to learn as many objects as possible.

Immediately after the study phase participants received the test phase with the stem competition task. It consisted of 76 stems. The stems appeared on a desktop screen with a constant speed of 10 seconds per each stimulus. After the 10 seconds the program moved to another stem. . This procedure was adapted using Jacoby's Process Dissociation Procedure

(Jacoby, 1991; Jacoby et al, 1993). The participants' responses were recorded on a blank sheet of paper by the experimenter.

Results

Main analysis

In order to achieve our objective that aimed to compare explicit and implicit memory performance across different learning environments we performed a Multivariate Analysis of Variance (MANOVA) with type of learning environment (a computerized measure, a 3D desktop environment and 3D virtual environment) as between factors.

MANOVA results using Pillai's trace point out a non-significant main effect of type of learning environment on the explicit and implicit memory performance, $V = 0.04$, $F(4, 148) = 0.82$, $p > .05$. This result reflects the absence of type of medium environment influence on explicit and implicit memory performance. That is, participants perform similar on explicit and implicit memory measures on a computerized measure, or on a 3D desktop environment or on a 3D virtual environment.

Supplementary analysis

Results from MANOVA point out no significant differences between the three learning environments on none of the cognitive absorption dimensions $V = 0.15$, $F(12, 140) = 0.97$, $p > .05$.

Further on, we have conducted correlational analysis between the perceived level of presence into the virtual environment and explicit and implicit memory performance. The association between explicit memory and presence was not statistically significant, $r(24) = -.02$, $p > .05$. Non-significant correlations have been obtained for implicit memory performance and total presence score, $r(24) = .15$, $p > .05$. Finally, the association between explicit and implicit memory for the 3D virtual environment is significant, $r(24) = -.47$, $p < .05$ with a medium magnitude of the effect size.

Discussion

The present research aimed to investigate the effect of three distinct learning environments on explicit and implicit memory performance. The learning environments contain different levels of immersion and task difficulty. Overall, results from the current research suggest that the type of learning environment does not influence explicit and implicit memory performance. Participants have similar performance on both controlled, conscious processes and uncontrolled, unconscious processes even if they learn on a computerized environment or on a 3D desktop environment with a low level of presence and task difficulty or on a 3D virtual environment with an increased level of presence and task difficulty. Further on, sense of presence into the virtual environment is not related with performance on neither explicit nor implicit memory. Finally, the current data bring support for the dual-process models of memory systems because controlled processes correlate negatively with uncontrolled processes.

Study 4. Do elder people learn different in a virtual environment? Proof of the concept.

Introduction

Considering the results from studies that investigated explicit and implicit memory in cognitive aging which point out a clear cognitive decline in older participants on explicit memory and mixed results for implicit memory in relation to age related deficits (Fleischman et al., 2004; Gopie et al., 2011; Graf, 1990; Howard et al., 1991; Jelicic et al., 1996; La Voie & Light, 1994; Light, & Singh, 1987) we aim to explore further the dissociation of memory process in older participants by using an ecological study task. We will apply a Process Dissociation procedure (Jacoby, 1991, 1993) to obtain estimates of explicit and implicit memory. Further on, we will compare the performance on explicit and implicit memory measures between two study/learning environments. One study environment is similar to a classical learning task in which participants are exposed to a series of pictures delivered on a desktop screen and have to learn them in a predefined period of time. The other study environment contains an analogue highly ecological task with identical stimuli embedded into a virtual apartment in which the participant has to perform the same learning task. An important aspect in the study of memory assessment using virtual reality is task difficulty. Overall, it is considered that assessment using virtual reality has an increased task difficulty expressed in poorer performance on cognitive tasks embedded in virtual reality (Adams et al., 2009; Neguț et al., in press). However, for memory assessment there are two opposite approaches. One states that memory performance in virtual reality is poorer because the replication of real world requires additional cognitive resources (Elkind et al., 2001; Gamberini, 2000). The opposite paradigm considers that memory performance is enhanced in virtual reality because a sense of presence, as well as visual realism boosts performance (Dinh, Walker, Hodges, Song, & Kobayashi, 1999; Lin, Duh, Parker, Abi-Rached, & Furness, 2002; Mania & Chalmers, 2001). Further on, cognitive absorption which is a state of deep involvement with software that predicts usage behavior (Agarwal & Karahanna, 2000) seems to be related to presence and immersion into virtual reality (Murray, Fox, & Pettifer, 2007).

In order to investigate explicit and implicit memory patterns in older participants under the influence of learning environment we aim to examine the following objectives:

1. To investigate differences between performance on explicit and implicit memory in two distinct learning environments, a classical measure and 3D virtual environment.

2. To compare the two environments on cognitive absorption.

Method

Participants and experimental design

The sample consisted of eight elderly participants aged between 60 and 77 years old (M = 66.87, SD = 6.05) out of which 5 (62.50%) were females with a drop-out rate of 46.66%. The participants were recruited from several social clubs for retirees across Cluj-Napoca City. They were highly functional without dementia and other neurological disorders.

We used a within-subject factorial experimental design with two experimental conditions. Each condition corresponds to one of the two distinct learning environments: a classical measure and 3D virtual environment. Participants were random assigned to the experimental conditions in the study phase: virtual reality testing first or second.

Measures

Sociodemographic variables. Participants' age and gender were reported.

Trail Making Test Part A and B (Reitan, 1958) was used as an executive function measure.

Simulator Sickness Questionnaire (SSQ, Kennedy, Lane, Berbaum, & Lilienthal, 1993) was used as a measure of Cybersickness.

Cognitive Absorption Scale (CAS, Agarwal & Karahanna, 2000) was adapted for the purpose of this research and used as a measure of cognitive absorption.

Presence was measured with Witmer and Singer Presence Questionnaire (PQ, Witmer & Singer, 1998).

Explicit and implicit memory measures were estimated using Process Dissociation Procedure (Jacoby, 1991; Jacoby et al., 1993). By applying an inclusion and exclusion test estimates of recollection and automatic or unconscious influences are obtained.

Materials

Virtual and computerized environment

In the present study we had two experimental conditions. For each condition we had to types of learning environment. The first environment corresponds to a classical neuropsychological assessment context in which the stimuli are delivered via a computer screen. The stimuli are delivered on a HP Z800 Workstation. The virtual environment that corresponds to the second learning environment runs on a CAVE Automatic Virtual Environment with four walls.

The virtual scenario consisted of a virtual apartment whit 3 rooms: a bedroom, an open space living room with a kitchen and a bathroom. It contained a total of 40 objects that are typical to an apartment (e.g. bed, chair, table, glasses, bookshelf, vase, and flower) (see Appendix 2). The computerized environment contained images of the same objects within the virtual environment. The images were randomly displayed on the computer desktop screen.

Procedure

The experiment had two main phases: study and test phase. The psychological scales related to technology and virtual reality exposer were administrated at the end of the experiment after completing the experimental procedure. In the study phase, the participants random assigned into the 3D desktop or 3D virtual reality condition were told that they will be guided through a virtual apartment. In the virtual apartment they will see different objects that are usually found into an apartment. As they will navigate into the apartment they will have to learn as many objects as possible.

Immediately after the study phase participants received the test phase with the stem competition task. It consisted of 76 stems. The stems appeared on a desktop screen with a constant speed of 10 seconds per each stimulus. After the 10 seconds the program moved to another stem. . This procedure was adapted using Jacoby's Process Dissociation Procedure (Jacoby, 1991; Jacoby et al, 1993). The participants' responses were recorded on a blank sheet of paper by the experimenter.

Results

In order to achieve our objective that aimed to compare explicit and implicit memory performance across different learning environments we performed the Wilcoxon signed-rank test with type of learning environment (a computerized measure and a 3D virtual environment) as within factors. Based on the small sample of participants we have decided that a more appropriate statistical test would be a non-parametric test that ranks the data, although a non-parametric test has a lower statistical power (Field, 2009). Results show that for explicit or controlled memory influences there were no statistically significant differences across conditions $z = -0.70$, $p > .05$, $r = .08$. Such a result suggests that participants perform similar on explicit memory despite the learning environment with a small magnitude of the effect size. For the implicit, automatic memory influences results suggest that the type of learning environment has no influence, $z = 0.00$, $p > .05$, $r = .00$.

Supplementary analysis

An analysis of simulator sickness showed that none of the participants reported Cybersickness as none had any severe symptoms. When we investigate the sense of presence into the virtual environment results point out that 50% of the participants reported medium to high levels of presence. Hence, only 2 participants reported a high level of presence and 2 participants presented lower level of presence.

Next we have compared differences between the computerized learning environment and the 3D virtual environment on the cognitive absorption. Results from the Wilcoxon signed-rank test point out that higher cognitive absorption for the virtual apartment compared to the computerized measure, $z = -2.04$, $p < .05$, with a small effect size, $r = .25$. One out of 8 participants had a higher score on cognitive absorption on the computerized measure while 7 out of 8 participants scored better on the virtual apartment condition.

Discussion

The main objective of this research paper was to address the effect of two distinct learning environments on explicit and implicit memory performance across a sample of elderly participants. We have chosen to investigate the effect of learning environment in case of older participants due to the fact that previous studies that focused on dual-memory processes in cognitive aging show a cognitive decline in older participants on explicit memory and mixed results for implicit memory in relation to age related deficits.

Because we had a small sample we have low statistical power, and we have slight chances to detect an effect. Therefore we have chances to commit type II error and to fail to detect an effect that is present. Our results should be interpreted with cautions in light of the proof of the concept principle.

As a concluding commentary, based on the data from the present research we might consider that type of learning environment yields a similar effect over explicit and implicit memory performance. There are no significant differences between the condition in which participants have to perform the learning task into a virtual environment and the condition in which they learn the same items on a computer desktop without immersion. However, participants seem to rate as more pleasant, enjoyable, and are more willing to use in the future the virtual environment than the computerized environment because higher scores were obtained in favor of virtual reality on the cognitive absorption dimension.

CHAPTER IV. GENERAL CONCLUSIONS AND DISCUSSION

Theoretical and practical implications

The findings from this Ph.D. thesis have several theoretical and practical implications that are related to directions in neuropsychological assessment. Overall, the results from the four studies conducted in this research paper bring evidence in favor of the validity of virtual reality-based measures. The findings provided by the studies clarify important issues concerning the task difficulty of the virtual reality measures and those related to the use of distractors as means to enhance ecological validity of the measures. Further on, the results from the studies can be interpreted as support for an increased level of ecological validity of the virtual reality-based measures.

Concerning validity issues the important information that the current Ph.D. thesis can provide is the fact that virtual reality-measures can be used as valid measures in neuropsychological assessment as they discriminate between different types of clinical populations such children with ADHD, brain injury, and schizophrenic patients. Also, they measure to some extent the same cognitive processes as classical and computerized measure.

We also have investigated the difficulty of cognitive tasks embedded in virtual reality. As mentioned earlier, the task difficulty of an assessment instrument is important in discriminating correctly patients from healthy participants. Also, if we look at task difficulty from an ecological point of view related to the similarity with the real world, we can conclude, based on the results that for executive functions assessment classical measures seem to under estimate the real performance that participants have in real life, because they perform better on classical measures compared to virtual reality. However, on explicit and implicit memory measures differences in performance between a virtual scenario and a computerized measure did not emerge. Adult and elderly participants have similar explicit memory and implicit memory performance despite the environment in which they performed the learning task. Therefore, we speculate that learning into the virtual environment has the same task difficulty as a classical computerized environment, targets the same cognitive processes and requires similar cognitive resources.

When it comes to the use of distractors for a more ecological assessment the empirically data obtained in this Ph.D. thesis offers mixed results to this assumption. Based on meta-analytic approach results show that distractors are not significant moderators. However, in Study 2 distractors influenced the attention performance in case of ADHD participants.

Another key finding of this Ph.D. thesis concerns the ecological validity of virtual reality-based measures. Previous research has speculated that neuropsychological assessment using virtual reality might have an increased level of ecological validity because it embeds cognitive tasks into a virtual environment that replicates accurately the everyday challenges. As a consequence, it is considered that virtual reality is more difficult. As presented earlier, current finding support this assumption of difficulty that can be considered as evidence for ecological validity. However, when we take into consideration the task difficulty of virtual reality measures applied on distinct cognitive processes, stronger evidence is gained for executive function and other cognitive measures, while for memory measures the evidence is limited. Results from the meta-analysis (Study 1) show a positive effect of virtual reality on memory measures. It seems that learning into a virtual environment boost memory performance. On Study 3 and 4 we failed to replicate this positive effect of virtual reality over memory performance as no differences emerge between virtual and computerized environments.

Limitations and future directions

Although results from the current research provide evidence for the validity of the virtual reality-based neuropsychological assessment future studies might consider to provide norms and to perform reliability analysis for virtual reality-based measures. Other studies can investigate the predictive validity of the virtual reality measures in relationship to real-life performance or other objective criteria. Future studies might consider upgrading the current graphics of virtual environments. The graphics of video games and movies are developing with fast speed and participants, especially children are usually connected to the new technological games releases. An upgraded version of the virtual environments' graphics might increase immersion, presence and enhance the similarity with the real world environment. Overall, almost half of the participants had reported medium to high levels of presence into the virtual environment. Also, another potential limitation concerns Cybersickness. Five healthy participants (Study 3) reported at least one severe symptom of simulator sickness. As suggested by the results Cybersickness correlated negatively with presence. We speculate that a diminished sense of presence into the virtual environment might yield negative effect over visual realism that facilitates the similarity with the real world. However, it appears that neither children, nor elderly participants had experienced Cybersickness. A possible explanation is the fact that children were immersed with a HMD while the healthy participants were immersed into the four walls ICube and navigated into the environment. Also, results show that in the same virtual environment delivered via an ICube with four walls elderly participants did not experience simulator sickness, while young adults did. We speculate that younger participants, are as children more connected with technological releases which might make them more demanding about the quality of the software graphics. The motion sickness can be enhanced by the navigation into the virtual environment. Future studies might consider testing this hypothesis because the simulator sickness might negatively influence presence and immersion. A major limitation concerns the Study's 4 low statistical power. Eight participants were enrolled in the study which is a small sample size. Therefore, we have chances to commit type II error and to fail to detect an effect that is present.

Finally, in case of Study 3 and 4 have limitation concerning control. Because we used a highly ecological environment such as the virtual apartment and asked participants to learn as many objects as possible while they are guided through the apartment and compared memory performance with an computerized environment that consisted of the same items from the apartment delivered item by item might have caused an non-equivalence regarding the time spent on encoding each item. However, because we wanted to target performance on a highly ecological assessment measure we can consider this potential threat to internal validity as an evidence for the ecological validity of the virtual environment-based assessment. Also, we did not ask participants to read aloud each object because it would be difficult to control for the speed each participant that learned into the virtual environment decided to allocate for encoding. We also did not control for the frequency and novelty of the items presented in the apartment.

First, almost half of the objects from the apartment were included in the original demo from EON Reality. However, the list of items that were added into the apartment was carefully inspected to exclude non-frequent and unusual items.

SELECTIVE REFERENCES

Adams, R., Finn, P., Moes, E., Flannery, K., & Rizzo, A. A. (2009). Distractibility in Attention/Deficit/ Hyperactivity Disorder (ADHD): the Virtual Reality Classroom. *Child Neuropsychology, 15*(2), 120-135. doi: 10.1080/09297040802169077Rizzo, A. 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. doi: 10.1080/09602010343000183

Anastasi, A., & Urbina, S. (1997). *Psychological testing*. Upper Saddle River, NJ: Prentice Hall.

Armstrong, C. M., Reger, G. M., Edwards, J., Rizzo, A. A., Courtney, C. G., & Parsons, T. D. (2013). Validity of the Virtual Reality Stroop Task (VRST) in active duty military. *Journal of Clinical and Experimental Neuropsychology, 35*(2), 113-123. doi: 10.1080/13803395.2012.740002

Baddeley, A., & Wilson, B. A. (1994). When implicit learning fails: Amnesia and the problem of error elimination. *Neuropsychologia*, *32*(1), 53–68. doi: 10.1016/0028-3932(94)90068-X

Barkley, R. A. (1991). The ecological validity of laboratory and analogue assessment methods of ADHD symptoms. *Journal of Abnormal Child Psychology, 19*(2), 149-178. doi: 10.1007/BF00909976

Bioulac, S., Lallemand, S., Rizzo, A., Philip, P., Fabrigoule, C., & Bouvard, M. P. (2012). Impact of time on task on ADHD patient's performances in a virtual classroom. *European Journal of Paediatric Neurology*, *16*, 514-521. doi: 10.1016/j.ejpn.2012.01.006

Elkind, J. S., Rubin, E., Rosenthal, S., Skoff, B., & Prather, P. (2001). A simulated reality scenario compared with the computerized Wisconsin Card Sorting Test: An analysis of preliminary results. *CyberPsychology & Behavior, 4*(4), 489-496. doi:

10.1089/109493101750527042

Gilboa, Y., Rosenblum, S., Fattal-Valevski, A., Toledano-Alhadef, H., Rizzo, A. S., & Josman, N. (2011). Using a Virtual Classroom environment to describe the attention deficits profile of children with Neurofibromatosis type 1. *Research in Developmental Disabilities, 32*(6), 2608–2613. doi: 10.1016/j.ridd.2011.06.014

Graf, P. (1990). Life-span changes in implicit and explicit memory. *Bulletin of the Psychonomic Society, 28*(4), 353-358. doi: 10.3758/BF03334041

Graf, P., & Schacter, D. L. (1985). Implicit and Explicit Memory for New Associations in Normal and Amnesic Subjects. *Journal of Experimental Psychology: Learning. Memory, and Cognition, 11*(3), 501-518. doi: 10.1037/0278-7393.11.3.501

Huang-Pollock, C. L., Karalunas, S. L., Tam, H., & Moore, A. N. (2012). Evaluating vigilance deficits in ADHD: a meta-analysis of CPT performance. *Journal of Abnormal Psychology, 121*(2), 360-371. doi: 10.1037/a0027205

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*(5), 513-541. doi: 10.1016/0749-596X(91)90025-F

Jacoby, L. L. (1996). Dissociating automatic and consciously controlled effects of study/test compatibility. *Journal of Memory and Language, 35*(1), 32-52. doi: 10.1006/jmla.1996.0002

Parsons, T. D., & Rizzo, A. A. (2008). Neuropsychological assessment of attentional processing using virtual reality. *Annual Review of CyberTherapy and Telemedicine, 6*(1), 23-28.

*Parsons, T. D., Bowerly, T., Buckwalter, J. G., & Rizzo, A. A. (2007). A controlled clinical comparison of attention performance in children with ADHD in a virtual reality classroom compared to standard neuropsychological methods. *Child Neuropsychology*, *13*(4), 363-381. doi: 10.1080/13825580600943473

Rizzo & Buckwalter, (1998). Virtual Reality and Cognitive Assessment and Rehabilitation: The State of the Art. In Giuseppe Riva (Ed.), *Virtual Reality in Neuro-Psycho-Physiology*. Amsterdam, Netherlands: Ios Press.

Rizzo, A. A., Bowerly, T., Buckwalter, J. G., Klimchuk, D., Mitura, R., & Parsons, T. D. (2006). A Virtual Reality Scenario for All Seasons: the Virtual Classroom. *CNS Spectrums*, *11*(1), 35-44. doi: 10.1017/S1092852900024196

Toth, J. P., Lindsay, D. S., & Jacoby, L. L. (1992). Awareness, automaticity, and memory dissociations. In L. R. Squire & N. Butters (Eds.), *Neuropsychology of Memory* (2nd ed., pp. 46-57). New York, NY: Guilford.

Urbina, S. (2004). *Essentials of psychological testing*. New Jersey: John Wiley & Sons.

Weniger, G., & Irle, E. (2008). Allocentric memory impaired and egocentric memory intact as assessed by virtual reality in recent-onset schizophrenia. *Schizophrenia Research*, *101*(1-3), 201-209. doi: 10.1016/j.schres.2008.01.011