Medicine Meets Virtual Reality 16
J.D. Westwood et al. (Eds.)
IOS Press, 2008

Virtual Environment for Assessment of Neurocognitive Functioning: Virtual Reality Cognitive Performance Assessment Test

Thomas D. PARSONS [a,1], Teri M. SILVA [a], Jarrell PAIR [a], and Albert A. RIZZO [a]
[a] *Institute for Creative Technologies, University of Southern California, Los Angeles*

Abstract. While standard neuropsychological measures have been found to have adequate predictive value, their ecological validity may diminish predictions about real world functioning. Virtual environments (VEs) are increasingly recognized as ecologically valid tools for neuropsychological assessment. We aim to develop a VE-based neuropsychological battery delivered within the context of a virtual city: Virtual Reality Cognitive Performance Assessment Test (VRCPAT). The 15 minute VRCPAT battery and 1.5 hour in-person neuropsychological assessment were conducted with a sample of 40 healthy adults, between the ages of 21 and 36, that included equivalent distributions of men and women from ethnically diverse populations. No subjects had history of psychiatric or neurologic conditions. Results supported both convergent and discriminant validity. The VRCPAT was found to correlate significantly with traditional neuropsychological tests assessing similar target constructs. No significant correlations existed between VRCPAT measures and non-target constructs. Findings suggest that the VRCPAT measures a capacity consistent with that of traditional neurocognitive measures; and is inconsistent with potential confounds. We conclude that the VRCPAT provides a unique opportunity to reliably and efficiently study neurocognitive function within an ecologically valid environment.

Keywords. Neuropsychological assessment, psychomtrics, ecological validity, virtual environment,

Introduction

Virtual environments (VE) are now being developed and validated that focus on component cognitive processes including: attention processes [1,2], spatial abilities [3-4], learning and memory [5-6], and executive functions [7]. The ability of VEs to create dynamic, immersive, three-dimensional stimulus environments, in which all behavioral responding can be recorded, offers assessment and rehabilitation options that are not available using traditional assessment methods [8-9]. The potential for increased ecological validity of neurocognitive batteries that include assessment using VEs may aid differential diagnosis and treatment planning. Much like aircraft simulators have been developed to assess and train piloting ability under a range of controllable stimulus conditions, at the University of Southern California's Institute for

[1] Corresponding Author: Thomas D. Parsons, Ph.D., Research Scientist and Neuropsychologist, University of Southern California, Institute for Creative Technologies, 13274 Fiji Way, Marina del Rey, CA. 90292-4019, E-mail: tparsons@usc.edu.

Creative Technologies we are developing VEs to establish a battery of tests that assess the specific neurocognitive components that underlie all facets of successful military performance: attention, spatial ability, memory, executive functioning and a host of higher-level language and reasoning abilities [10]. We aim to develop a VE-based neuropsychological battery delivered within the context of a Middle Eastern city VE scenario: Virtual Reality Cognitive Performance Assessment Test (VRCPAT). After a brief discussion of our VE-based neuropsychological battery, we report on the psychometric properties of data gained from human pilot testing.

1. Virtual Reality Cognitive Performance Assessment Test (VRCPAT)

The VRCPAT project focuses on the refined analysis of neurocognitive testing using a VE to assess recall of targets delivered within the context of a virtual city. The VRCPAT is a three-dimensional virtual Iraqi city environment that was designed to run on a Pentium IV notebook computer with one gigabyte RAM and a 128 megabyte graphics card. The primary aim of the VRCPAT project is to use the already existing library of assets as the basis for creating a VE for the standardized assessment of neurocognitive performance within a contextually relevant VE. The application uses USC's FlatWorld Simulation Control Architecture (FSCA). The FSCA enables a network-centric system of client displays driven by a single controller application. The controller application broadcasts user-triggered or scripted-event data to the display client. The real-time three-dimensional scenes are presented using Numerical Design Limited's (NDL's) Gamebryo graphics engine. The content was edited and exported to the engine, using Alias's Maya software. Three-dimensional visual imagery is presented using the eMagin z800. Navigation through the scenario uses a common USB Logitech game pad device.

Virtual reality-based simulation technology approaches, as delineated herein, are considered to be the future alternative for devising neurocognitive assessment measures that will have better ecological/predictive validity for real-world performance. As well, the flexibility of stimulus delivery and response capture that are fundamental characteristics of such digital environments is viewed as a way for research objectives to be addressed in a more efficient fashion for long term needs. In this regard, while the structure of the VRCPAT assessment task appears deceptively simple, the overall design of this type of assessment tool allows for tremendous flexibility in terms of the independent variables that could be studied with this method once the psychometric properties of the standardized test are determined. Such flexibility would allow for this system to be viewed as an open platform on which a wide range of research questions. These include the manipulation of: 1) information load on the front end via the complexity of target stimuli to be attended to and the type of information in terms of relevance, complexity, similarity, vagueness, sensory properties, etc.; 2) temporal constraints during varied sustained assessment conditions; 3) distracting activities during the neurocognitive assessments; 4) sensory modality of the information presentation that needs to be attended to; 5) the reward structure used during some tests to assess motivational factors that influence performance; 6) the presentation of aversive stimuli for stressed performance evaluations; and 7) the development of a testbed whereby cognitive training and augmented cognition strategies could be tested under known conditions supported by normative standards.

2. Psychometric Validation

Following the general psychometric conventions, any measure purporting to quantify a particular neurocognitive domain should be highly correlated with other measures of the same neurocognitive domain (convergent validity), whereas it should not be too highly correlated with tests of different neurocognitive domains (discriminant validity). Herein we report on the psychometric properties of data gained from human pilot testing with the VRCPAT.

We acquired data on the implementation of VRCPAT in a normative sample that also received a traditional paper and pencil battery. Because the VRCPAT was designed to tap very specific neurocognitive systems and not to mirror a traditional paper-and-pencil battery, our goal is not to replace the traditional battery for all neurocognitive domains. We aim to assess the psychometric properties of the VE and paper-and-pencil measures. Hence, scores were correlated with demographic and other performance tests measures administered. Standard correlational analyses using a brief demographic survey and pencil-and-paper cognitive tests aid our initial assessment of both the convergent and divergent validity properties of this form of assessment.

Our plan for the development and implementation of the VRCPAT's psychometric properties involved systematic refinement analyses that acted as a component of an ongoing dialectic between measurement and substantive research. We aim to make the VRCPAT a well developed measure that facilitates substantive advances. We determined the content homogeneity of each of the VRCPAT's unidimensional facets. The establishment of the VRCPAT's psychometric properties removed the possibility that results reflect correlates of the target construct (memory and/or attention) but are not prototypic of it. We also assessed the level to which all aspects of the target construct (memory) is under- or overrepresented in the VRCPAT's composition, and assess whether the experience of some aspects of the virtual environment introduced variance unrelated to the target construct.

2.1 Participants

The study sample included 40 healthy subjects (Age, mean = 24.45, SD = 3.05; 50 % male; and Education, mean = 14.05, SD = 0.51). Strict exclusion criteria were enforced so as to minimize the possible confounding effects of comorbid factors known to adversely impact cognition, including psychiatric (e.g., mental retardation, psychotic disorders, diagnosed learning disabilities, Attention-Deficit/Hyperactivity Disorder, and Bipolar Disorders, as well as substance-related disorders within two years of evaluation) and neurologic (e.g., seizure disorders, closed head injuries with loss of consciousness greater than 15 minutes, and neoplastic diseases) conditions. Subjects were comparable in age, education, ethnicity, sex, and self-reported symptoms of depression.

2.2 Procedure

The University of Southern California's Institutional Review Board approved the study. Experimental sessions took place over a two hour period. After informed consent was obtained, basic demographic information and computer experience and usage activities were recorded. Subjects then completed a neuropsychological battery administered

under standard conditions. Following completion of the neuropsychological battery, subjects completed the simulator sickness questionnaire [11], which includes a pre-VR exposure symptom checklist. Next, all participants were administered the VRCPAT as part of the larger neuropsychological test battery.

The following traditionally used paper and pencil neuropsychological measures were used as convergent validity measures, because each is considered to have an important memory component and has been used clinically to estimate memory abilities: To assess verbal learning and memory we used the Hopkins Verbal Learning Test – Revised; to assess nonverbal learning and memory we used the Brief Visuospatial Memory Test – Revised; and to assess Lexical-Semantic Memory we used Controlled Oral Word Association Test; 2) Semantic Fluency (Animals).

Discriminant validity measures that were drawn from the corpus of traditionally used paper and pencil neuropsychological measures included: to assess Attention we used Digit Span (Forward and Backward) from the Wechsler Adult Intelligence Scale –Third edition (WAIS-III); to assess processing speed we used Digit Symbol Coding from the Wechsler Adult Intelligence Scale –Third edition (WAIS-III), and Trail Making Test Part A (TMT-A); to assess executive functioning we used Trail Making Test Part B (TMT-B) and the Stroop Color and Word Test.

The VRCPAT is a 15-minute measure, in which participants (referred to as "users" in the following text) then go through the following steps: First, users are presented with 10 pieces of language-based information to be learned, without any context for what they will need to do with this information. The acquisition phase is initially standardized to three one-minute trials. At the end of each trial, users are then asked to name the objects that they studied as an assessment of initial declarative recall memory. After users are given the three one-minute trials to "memorize" the stimuli, a brief "interface training" period then occurs in which users become familiar with their objective, the controls of the game pad navigation interface and head-mounted display (HMD). The task is read aloud by the investigator and contains specific instructions for how to proceed through the VE and how to record images of each target object. Once users indicate that they are comfortable within the VE and can demonstrate comprehension of the navigation interface and targeting procedure, the investigator asks if there are any questions. If so, clarification and coaching occur until the user can fully comprehend the task. After completing the VE testing, users are asked to recall the original list of stimuli and at which target zones they were found. A trained research assistant administered all psychometric tests. The Simulator Sickness Questionnaire (SSQ; [11]) was used to determine whether the participant felt sick as a result of the VR experience.

3. Results

Given the similarity of participants in terms of age, sex, education, immersiveness, and ethnicity, no correction for these variables was employed. Notably, none of the participants reported simulator sickness following VR exposure as measured by the SSQ. To provide preliminary data to support the validity of the VRCPAT as a measure of learning and memory, recall indices from the VRCPAT and traditional neuropsychological tests were correlated. Indices were developed from linear composites derived from z-score transformations. Specifically, Pearson correlation

analyses were used to compare recall from the VRCPAT with linear composites derived from traditional neuropsychological measures.

3.1 Convergent Validity Tests

Whilst the VRCPAT Total Memory Score was significantly correlated with composites derived from established measures of learning and memory, it did not correlate with possibly confounded variables (i.e., Executive Functions; Attention; and Processing Speed) drawn from traditional neuropsychological measures that are not assessments of learning and memory. Hence, the results indicated that the VRCPAT correlated significantly with the traditional neuropsychological Learning Composite (HVLT Trials 1-3; and BVMT Trials 1-3; $r = 0.68$, $p < 0.001$), with 46% variance shared between the two indices. The results indicated that the VRCPAT also correlated significantly with the traditional neuropsychological Memory Composite (HVLT Total Recall after a Delay; and BVMT Total Recall after a Delay; $r = 0.67$, $p < 0.001$), with 45% variance shared between the two indices.

3.2 Discriminant Validity Tests

As expected, there were no significant correlations between VRCPAT measures and the following neuropsychology test composites: Executive Functions Composite; Attention Composite; or Processing Speed Composite. Hence, each of the discriminant validity significance tests were as predicted, that is, did not correlate with theoretically unrelated abilities. Although validity coefficients drawn from composites may not meet validity expectations it may still be the case that individual measures account for some of the trait variance. Therefore, we assessed the measures both as composites and individually. As such, we compared the VRCPAT with the actual neuropsychological tests (used to derive the Learning Composite and the Memory Composite). Analysis of the relations between the VRCPAT Total Memory Score and the actual learning and memory tests revealed significant correlations for each of the convergent validity significance tests, in accordance with prediction. For correlations between the VRCPAT and traditional psychometric measures we only considered those correlations that met the criterion of $p < .05$ to be meaningful. Given our small sample size we kept P at this level, despite the risk of Type I error with multiple correlations. All of our significant correlations were associated with at least moderate effect sizes.

4. Discussion

The results of this study indicate that: 1) VRCPAT memory measures correlated significantly with scores from the memory measures drawn from the traditional neuropsychological test battery; 2) VRCPAT memory scores did not correlate with non-memory measures drawn from the traditional neuropsychological test battery. Additionally, no negative side effects were associated with use of the VRCPAT. The establishment that the VRCPAT's memory measures correlated significantly with scores from the memory measures drawn from the traditional neuropsychological test battery but not with non-memory measures removed the possibility that results

reflected correlates of the non-target construct (i.e. processing speed; executive function).

Our goal was to conduct an initial pilot study of a VRCPAT that employs a standard neuropsychological battery for the assessment of normal participants. We believe that this goal was met. We recognize, however, that the current findings are only a first step in the development this tool. Many more steps need to be taken in order to continue the process of test development and to fully establish the VRCPAT as a measure that contributes to existing assessment procedures for the diagnosis of memory decline. Whilst the VRCPAT as a measure needs to be fully validated, current findings provide preliminary data regarding the validity of the virtual environment as a memory measure. The VRCPAT was correlated with widely used memory assessment tools. Nevertheless, the fairly small sample size requires that the reliability and validity of the VRCPAT be established using a larger sample of well-matched participants. This will ensure that current findings are not a sample size related anomaly. Finally, the ability of the VRCPAT to accurately classify participants not involved in the initial validation study will need to be examined for cross-validation purposes.

References

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

[2] Rizzo, A.A, Bowerly, T., Buckwalter, J., Klimchuk, D., Mitura, R., Parsons, T.D. A Virtual Reality Scenario for All Seasons: The Virtual Classroom. *CNS Spectrums* **11** (2006), 35-44.

[3] Parsons, T.D., Larson, P., Kratz, K., Thiebaux, M., Bluestein, B., & Buckwalter, J.G. Sex differences in mental rotation and spatial rotation in a virtual environment. *Neuropsychologia* **42** (2004), 555–562.

[4] Parsons, T.D., Rizzo, A.A., van der Zaag, C., McGee, J.S., & Buckwalter, J.G. Gender and Cognitive Performance: A Test of the Common Cause Hypothesis. *Aging, Neuropsychology, and Cognition* **12** (2005), 78-88.

[5] Parsons, T.D., & Rizzo, A.A. Initial Validation of a Virtual Environment for Assessment of Memory Functioning: Virtual Reality Cognitive Performance Assessment Test. *Cyberpsychology and Behavior* (in press).

[6] Matheis, R.J., Schultheis, M.T., Tiersky, L.A., Deluca, J., Millis, S.R., & Rizzo, A. Is learning and memory different in a virtual environment? *The Clinical Neuropsychologist* **21** (2007), 146-161.

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

[8] Rizzo, A.A., Strickland, D., & Bouchard, S. The challenge of using virtual reality in telerehabilitation. *Telemedicine Journal and E-Health* **10** (2004), 184-95.

[9] Rizzo, A.A., & Buckwalter, J.G. Virtual reality and cognitive assessment and rehabilitation: the state of the art. *Studies in Health Technology and Informatics* **44** (1997), 123-45.

[10] Parsons, T. D., Rizzo, A. A., Bamattre, J., & Brennan, J. Virtual reality cognitive performance assessment test. *Annual Review of CyberTherapy and Telemedicine* **5** (2007), 163-171.

[11] Kennedy, R. S., Fowlkes, J. E., Berbaum, K. S., & Lilienthal, M. G. Use of a motion sickness history questionnaire for prediction of simulator sickness. *Aviation, Space, and Environmental Medicine* **63** (1992), 588–593.