Initial Validation of a Virtual Environment for Assessment of Memory Functioning: Virtual Reality Cognitive Performance Assessment Test

THOMAS D. PARSONS, Ph.D. and ALBERT A. RIZZO, Ph.D.

ABSTRACT

The current project is an initial attempt at validating the Virtual Reality Cognitive Performance Assessment Test (VRCPAT), a virtual environment–based measure of learning and memory. To examine convergent and discriminant validity, a multitrait–multimethod matrix was used in which we hypothesized that the VRCPAT's total learning and memory scores would correlate with other neuropsychological measures involving learning and memory but not with measures involving potential confounds (i.e., executive functions; attention; processing speed; and verbal fluency). Using a sequential hierarchical strategy, each stage of test development did not proceed until specified criteria were met. The 15-minute VRCPAT battery and a 1.5-hour in-person neuropsychological assessment were conducted with a sample of 30 healthy adults, between the ages of 21 and 36, that included equivalent distributions of men and women from ethnically diverse populations. Results supported both convergent and discriminant validity. That is, findings suggest that the VRCPAT measures a capacity that is (a) consistent with that assessed by traditional paper-and-pencil measures involving learning and memory and (b) inconsistent with that assessed by traditional paper-and-pencil measures assessing neurocognitive domains traditionally assumed to be other than learning and memory. We conclude that the VRCPAT is a valid test that provides a unique opportunity to reliably and efficiently study memory function within an ecologically valid environment.

INTRODUCTION

MEMORY ASSESSMENT is one of the principal objectives of neuropsychological evaluation. Memory is generally accepted by neuropsychologists to be part of a variety of brain processes that involve the capacity to encode, store, and retrieve information.[1] Given the fundamental functions involved in memory and its impact on activities of daily living, memory researchers seek to understand the neural correlates of memory.[2] Additionally, the brain's examination of material stored in memory requires that the neural connections process observations at multifarious levels in a connected manner.[3–4] Emerging neurobiological views of hippocampal function make efforts to include other levels of analysis: anatomy, physiology, and plasticity.[5] While standard neuropsychological measures have been found to have adequate predictive value, their ecological validity may diminish predictions about real-world functioning.[6–9] Traditional neurocognitive measures may not replicate the diverse environment in which persons live. Additionally, standard neurocognitive batteries tend to examine isolated components of neuropsychological ability, which may not accurately reflect distinct cognitive domains.[10–11]

Virtual reality (VR) applications that focus on component cognitive processes, including attention processes,[12] spatial abilities,[13–14] memory,[15] and executive functions,[16–17] are now being developed and tested. The increased ecological validity of neurocognitive batteries that include assessment using VR scenarios may aid differential diagnosis and treatment planning.

Within a virtual environment (VE), it is possible to systematically present cognitive tasks targeting memory performance beyond what are currently available using traditional methods.[15,18–20] Reliability of memory assessment can be enhanced in VR by better control of the perceptual environment, more consistent stimulus presentation, and more precise and accurate scoring. VEs may also improve on the validity measurement via the quantification of more discrete behavioral responses, allowing for the identification of more specific cognitive domains. VR could allow for memory to be tested in situations that are more ecologically valid. Participants can be evaluated in an environment that simulates the real world, not a contrived testing environment.

Research examining memory for objects and the spatial layout contained in a VE has been underway since the mid-1990s at the University of East London (UEL). The UEL group has focused on specifying the types of memory that may be assessed during a four-room house navigation task with an aim toward targeting memory deficits in clinical populations.[21–23] This approach used a non-head-mounted display flat-screen system with a joystick interface, which allowed one participant to navigate the house (active condition) while a yoked participant was simply exposed to the same journey but had no navigational control (passive condition). Both participants were directed to seek out an object (e.g., toy car) during the exploration, and differential memory performance between the two groups on spatial versus object memory of the environment was tested. In initial tests with normal populations, it was observed that the active groups showed better spatial memory for the route, while the passive group displayed superior object recall and recognition memory for the items viewed along the route.[21] These results were then replicated by collaborators with this group.[23] Mixed results have been reported by other non-head-mounted display VE studies in this area, one showing active participation enhancing spatial memory[25] and no differences reported in the others.[26–27] However, the spatial orientation tasks used in these latter studies differed in the degree to which subjects were allowed to "retrace their route," and this may have

mitigated the procedural training component's contribution to the observed outcomes (no difference between active vs. passive exposure).[23] Since that time, others have done similar research on spatial navigation and memory but have yet to tackle the challenge for creating a norm-based approach within a standardized VE.

This research was expanded using the Rose/Attree scenario with clinical populations (i.e., multiple sclerosis and stroke) and produced results that support the value of this sort of VE application to inform assessment needs. In recent studies using this VR application, as expected, stroke patients performed statistically less well than did unimpaired participants. What is more interesting, though, is that while the typical spatial–content memory dissociation was found with unimpaired groups (i.e., active = better spatial memory; passive = better object memory), and the active stroke group displayed better spatial memory than the passive group, the stroke patients displayed no advantage on object memory while in the passive condition.[23] Similar findings using this scenario were also reported in a study[24] comparing multiple sclerosis (MS) patients with a healthy group. These findings also could not be accounted for via relationships with standard pencil-and-paper memory-assessment tools. This type of impairment in explicit incidental memory observed in patients with MS suggests that a VR approach may be of particular value in detecting subtle deficits in these patients.[24] This is in line with previous work[28] in which participants recalled half as much in a VE that contained the sort of distracters present in everyday life when compared to "static" presentations, as are typically used in standard psychometric methods. From this finding, the targeting of memory processes within a VE may provide more unique and specific information that could enhance our understanding of normal cognitive processes as well as underscore the potential value of immersive approaches that promote "procedural" involvement in the design of better memory assessment approaches needed for predicting real-world performance.

The Virtual Reality Cognitive Performance Assessment Test (VRCPAT) builds on these projects and focuses on refined analysis of neurocognitive testing using a VE to assess recall of targets delivered within the context of a virtual city. Specifically, the primary aim of the present study was to examine the construct validity of the VRCPAT using the methodology provided by the multitrait–multimethod matrix.[29] The use of this matrix approach with multiple neurocognitive measures allows the simultaneous investigation of convergent validity

(i.e., extent to which different neurocognitive measures of memory are related) and discriminant validity (i.e., extent to which neurocognitive measures of domains other than memory are unrelated). The use of the multitrait–multimethod matrix gave us the advantage of being able to examine method variance (i.e., degree to which scales are correlated because they use the same method of measurement rather than because they share valid trait variance). Using "pure" measures of memory introduces less noise into the decision matrix.[30] Following the multitrait–multimethod matrix, 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).

In our assessment of convergent validity, we hypothesized that the performance on the VRCPAT would correlate significantly with performance on traditional neuropsychological measures of memory. Given the enhanced ecological validity of VR measures, the VRCPAT was assumed to provide a more pure measure of memory than traditional neuropsychological tests. As a result, we hypothesized that the correlations would be moderate rather than high. In our assessment of discriminant validity, we hypothesized that correlations between the VRCPAT and traditional neuropsychological measures of domains other than memory would not be statistically significant.

METHODS

We acquired data on the implementation of a VE (i.e., VRCPAT) in a normative sample that also received a traditional paper-and-pencil battery. We aimed to assess the psychometric properties of the VR 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 aided our initial assessment of both the concurrent and divergent validity properties of this form of assessment.

Participants

The study sample included 30 healthy participants (15 males and 15 females) between the ages of 21 and 36 ($M = 24.97$, $SD = 3.78$). Participants were initially recruited from undergraduate and graduate schools; education levels ranged from 13 to 20 years ($M = 16.13$, $SD = 1.69$). Ethnicity was as follows: Caucasian ($n = 17$), African American ($n = 2$), Hispanic ($n = 3$), Native American ($n = 1$), and Asian Pacific ($n = 7$). Participants were comparable in age, education, ethnicity, sex, and self-reported symptoms of depression.

Inclusion and exclusion criteria

Strict exclusion criteria were enforced to minimize possible confounding effects of comorbid factors known to adversely impact cognition, including psychiatric conditions (e.g., mental retardation, psychotic disorders, diagnosed learning disabilities, attention deficit/hyperactivity disorder, and bipolar disorders, as well as substance-related disorders within 2 years of evaluation) and neurologic conditions (e.g., seizure disorders, closed head injuries with loss of consciousness greater than 15 minutes, and neoplastic diseases).

Apparatus

The VRCPAT, a 3D virtual city environment, was designed to run on a Pentium 4 notebook computer with 1 GB RAM and a 128-MB DirectX 9 compatible graphics card. The primary aim of the current project was to use the already existing library (University of Southern California, Institute for Creative Technologies) of assets as the basis for creating a VR application for the standardized assessment of memory performance within a contextually relevant VE.

The application uses the Institute for Creative Technologies' 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 3D 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. We believe that the head-mounted display (HMD) approach provides the optimal level of immersion and interaction for this application at a cost that now rivals that of a high-quality flat-screen display. The VRCPAT software is a 3D VE programmed to simulate a city environment. During immersion in the VRCPAT, participants are seated at a desk and have a complete 360-degree view of the city environment.

Procedure

The University of Southern California's Institutional Review Board approved the study. Experimental sessions took place over a 2-hour period. After informed consent was obtained, basic demographic information and computer experience and usage activities were recorded. Participants completed a neuropsychological battery administered under standard conditions, then completed the simulator sickness questionnaire,[31] which includes a pre-VR exposure symptom checklist. Next, all participants were administered the VRCPAT as part of a larger neuropsychological test battery.

Traditional neuropsychological battery. The following 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: the Hopkins Verbal Learning Test—Revised (HVLT-R)[32] and the Brief Visuospatial Memory Test—Revised (BVMT-R).[33] The following were used as discriminant validity measures: (a) Controlled Oral Word Association Test (FAS);[34] (b) Semantic Fluency (Animals);[35] (c) Digit Symbol Coding and Digit Span (Forward and Backward) from the Wechsler Adult Intelligence Scale, Third Edition (WAIS-III);[36] (d) Trail Making Test Parts A and B (TMT);[37–38] (e) Stroop Color and Word Test;[39] and (f) Wechsler Test of Adult Reading (WTAR).[36]

The VRCPAT is a 15-minute measure in which participants (or "users") completed the following steps:

Acquisition phase (VRCPAT's learning score). Users were presented with 10 pieces of language-based information to be learned without any context for what they would need to do with this information. The information stimuli was primarily language based (blue car, intact barrel, red shipping container with numbers 7668, etc.), although stimuli included minimal imagery to provide opportunities for more context-relevant stimulus creation (e.g., a crate with a flag in an upper corner [with just an image of a flag presented]). The acquisition phase was initially standardized to three 1-minute trials. At the end of each trial, users were asked to name the objects they studied as an assessment of initial declarative recall memory. A psychometrist scored the VRCPAT's learning score by recording the total recall (sum of Trials 1 through 3).

VR interface and task training phase. Next, a brief "interface training" period occurred in which users became familiar with their objective, the controls of the game pad navigation interface, and the HMD. The task was read aloud by the investigator and contained specific instructions for how to proceed through the VE and how to record images of each target object. Users were given as much time as needed to explore a limited area of the environment. This exploration area was determined by the experimenter. During this phase, the investigator could informally present verbal guidance to the users to help them become familiar with the interface navigation, response button, and HMD viewing parameters. This phase was designed to teach the interface controls to the users so that performance on the VE navigation and object selection interaction tasks would not distract users in the retrieval phase.

Retrieval phase (VRCPAT's memory score). Once users indicated they were comfortable within the VE and demonstrated comprehension of the navigation interface and targeting procedure, the investigator asked if there were any questions. If so, clarification and coaching occurred until the users fully comprehended the task. Next, the users "traveled" to five zones in the city. At each of the five zones, the users were exposed to both targets (i.e., items from the 10-item list in the acquisition phase) and foils (i.e., items that were similar to or different from—but not identical to—the targets). A psychometrist scored the participants' performance during the acquisition phase, which made up part of the VRCPAT's memory score. The other part of the VRCPAT memory score resulted from the debriefing phase. The psychometrist recorded the total delayed recall from within the five zones. Each of the five zones had two targets. Each target was scored as 1 point per recall for a total of 10 possible points.

Debriefing phase. During this phase, users were asked to recall the original list of stimuli and at which target zones they were found. The performance measures derived from this test included number of correct hits, false hits, time to successfully complete per target zone, and time to complete overall. A psychometrist scored the VRCPAT's memory score by recording the total delayed recall both from within the five zones and again at the end of the VR scenario. In the debriefing phase, the participant would recall each target (10 possible points) and the zone in which the target was found (10 possible points). The Simulator Sickness Questionnaire (SSQ)[31] was used to determine whether the participant felt sick as a result of the VR experience.

Data analytics

All data were analyzed using SAS version 9.1. Descriptive statistics were calculated for participant demographics and for results of the VRCPAT and the criterion neuropsychological tests. Missing data were imputed by either mean substitution or last case carried forward. Two types of analyses were performed. The first involved computing basic correlations between VRCPAT measures and traditional neuropsychological measures assessing learning and memory. Two-tailed Pearson correlations were used. To make the results for continuous tests comparable, the test scores were converted to standard scores with a mean of zero and standard deviation of one (*z*-score transformation). The second series of analyses involved computing basic correlations between VRCPAT measures and traditional neuropsychological measures assessing domains traditionally understood to be other than learning and memory. Two-tailed Pearson correlations were used. To make the results for continuous tests comparable, the test scores were converted to standard scores with a mean of zero and standard deviation of one (*z*-score transformation). Following the multitrait–multimethod matrix, we assessed whether the monotrait–heteromethod coefficients were higher than their corresponding heterotrait–heteromethod coefficients. In other words, efforts to measure VRCPAT memory by standard neuropsychological memory measures should yield higher correlations than efforts to measure VRCPAT memory by standard neuropsychological measures of domains other than memory.

RESULTS

Given the similarity of participants (in terms of performances for both the VRCPAT and the standard neuropsychological measures) on age, sex, education, ethnicity, and simulator experience (as measured by the SSQ), no correction for these variables was employed. Table 1 presents means and standard deviations of all variables.

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.

Convergent validity tests

While the VRCPAT total memory score was significantly correlated (as expected) with composites derived from established measures of learning and memory, it did not correlate (again, as expected) with traditional neuropsychological domain composites (see Table 2). The results indicated that the

TABLE 1.   MEANS AND STANDARD DEVIATIONS (RAW SCORES)

	Mean	Variance	Standard deviation	Standard error
VRCPAT				
Total learning	35.67	189.06	13.75	2.51
Total memory	6.04	5.05	2.25	0.41
Neuropsychological measures				
HVLT Trials 1–3	27.90	21.40	4.63	0.84
BVMT Trials 1–3	28.90	20.51	4.53	0.83
HVLT Retention	10.27	3.65	1.91	0.35
BVMT-R Retention	11.00	2.21	1.49	0.27
Trail Making Test, Part B	46.30	184.70	13.59	2.48
Stroop Interference	97.07	373.44	19.32	3.53
Digit Span Forward	7.37	1.34	1.16	0.21
Digit Span Backward	6.00	2.21	1.49	0.27
Trail Making Test, Part A	24.10	82.02	9.06	1.65
Digit-Symbol Coding	86.10	310.16	17.61	3.22
Semantic Fluency	27.20	29.20	5.40	0.99
Phonemic Fluency	47.97	139.14	11.80	2.15

Note: For all analyses, *N* = 30.
HVLT, Hopkins Verbal Learning Test; BVMT-R, Brief Visuospatial Memory Test—Revised.

TABLE 2. CORRELATIONS BETWEEN TRADITIONAL TEST COGNITIVE DOMAIN SCORES AND VRCPAT LEARNING AND MEMORY SCORES

	VRCPAT learning		VRCPAT memory	
	r	p	r	p
Learning	0.69	<0.01	0.73	<0.01
Memory	0.67	<0.01	0.65	<0.01
Executive functions	0.21	0.25	0.13	0.48
Attention	0.25	0.18	0.12	0.50
Processing speed	0.12	0.50	0.15	0.43
Verbal fluency	0.27	0.15	0.25	0.18

Note: For all analyses, $N = 30$.

The composites were composed as follows: learning: Trials 1–3 of both the Hopkins Verbal Learning Test—Revised (HVLT-R) and Brief Visuospatial Memory Test—Revised (BVMT-R); memory: delayed recall from the HVLT and BVMT-R; executive functions: Stroop Color and Word Test, and Trail Making Test, Part B; attention: Digit Span (Forward and Backward) from the Wechsler Adult Intelligence Scale, 3rd edition; processing speed: Trail Making Test, Part A and Digit Symbol Coding from the Wechsler Adult Intelligence Scale, 3rd edition; verbal fluency: FAS and Animals from the Controlled Oral Word Association Test.

"r," Pearson correlation; "p," probability.

*$p < 0.01$.

VRCPAT correlated significantly with the traditional neuropsychological learning composite (HVLT Trials 1–3 and BVMT Trials 1–3; $r = 0.69, p < 0.001$), with 48% of the variance shared between the two indices. The results indicated that the VRCPAT 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% of the variance shared between the two indices.

Given that even the validity coefficients exceeding the heterotrait–monomethod correlations have been suggested to still be potentially useful in accounting for some of the trait variance if combined with other measures, 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 four convergent validity significance tests, in accordance with prediction (Table 3).

Discriminant validity tests

There were no significant correlations between VRCPAT measures and the following neuropsychology test composites: executive functions composite, attention composite, processing speed composite, and verbal fluency composite. Hence, each of the four discriminant validity significance tests were as predicted, that is, did not correlate with theoretically unrelated abilities (Table 2).

As in the convergent validity assessment, validity coefficients exceeding the heterotrait–monomethod correlations may still be potentially useful in accounting for some of the trait variance if combined with other measures. Hence, we assessed the measures both as composites and individually. As such, we compared the VRCPAT with the actual neuropsychological tests (used to derive the execu-

TABLE 3. CONVERGENT VALIDITY: CORRELATIONS BETWEEN TRADITIONAL NEUROCOGNITIVE TEST SCORES AND VRCPAT TOTAL MEMORY SCORE

	r (x,y)	r^2	t	p
Learning				
HVLT Trials 1–3	0.58	0.34	3.78	<0.01
BVMT-R Trials 1–3	0.75	0.56	6.01	<0.01
Memory				
HVLT Retention	0.50	0.25	3.05	<0.01
BVMT-R Retention	0.40	0.16	2.29	<0.01

Note: For all analyses, $N = 30$.

HVLT, Hopkins Verbal Learning Test; BVMT-R, Brief Visuospatial Memory Test—Revised. "r(x,y)" Pearson correlation between the independent variable and the dependent variable; "r^2," coefficient of determination; "T" student's t-test; "p" probability.

tive function composite, attention composite, processing speed composite, and verbal fluency composite). Analysis of the relations between the VRCPAT total memory score and the paper-and-pencil measures that are not traditionally believed to reflect learning and memory (executive function, attention, processing speed, and verbal fluency) tests did not reveal significant correlations for any of the discriminant validity significance tests, in accordance with prediction (Table 4).

For correlations between the VRCPAT and traditional psychometric measures we only considered those correlations that met the criterion of $p < 0.05$ to be meaningful. Given our small sample size, we kept alpha 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.

DISCUSSION

This study provides preliminary validation of the VRCPAT's memory module. Convergent and discriminant validity were evaluated using neuropsychological tests chosen a priori, according to the multitrait–multimethod matrix approach. The VRCPAT was significantly related to all of the other measures (both individual and composites) used to assess learning memory, in accordance with prediction. Further, following expectation, VRCPAT memory scores did not correlate with nonmemory measures drawn from the traditional neuropsycho-

logical test battery. Together, these findings suggest that the VRCPAT assesses a construct that is similar to those measured by the other memory tests in this study.

The use of the multitrait–multimethod analyses allowed us to examine the extent of construct validity. Accordingly, we concluded that the VRCPAT had appropriate levels of convergent and divergent validity in that the degree to which convergent validity coefficients (assessing memory domain) derived from the VRCPAT memory score and the traditional neuropsychological measures of memory were larger than correlations of different measures assessing domains other than memory within the same array of measures. Evidence for discriminant validity was indicated in that correlations of different scales assessed using different measures were lower than the convergent validity coefficients.

The establishment that the VRCPAT's memory measure correlates significantly with scores from memory measures drawn from the traditional neuropsychological test battery, but not with nonmemory measures, removed the possibility that results reflected correlates of the nontarget construct (i.e., executive function, attention, processing speed, and verbal fluency).

Our findings should be understood in the context of some limitations. These findings are based on a fairly small sample size. As a necessary next step, the reliability and validity of the test needs to be established using a larger sample of participants to ensure that the current findings are not an anomaly due to sample size. Additionally, as indicated pre-

TABLE 4. DISCRIMINANT VALIDITY: CORRELATIONS BETWEEN TRADITIONAL
NEUROCOGNITIVE TEST SCORES AND VRCPAT TOTAL MEMORY SCORE

	$r\ (x,y)$	r^2	t	p
Executive functions				
Trail Making Test: Part B	0.34	0.12	1.94	0.07
Stroop Interference	0.30	0.09	1.66	0.11
Attention				
Digit Span Forward	0.13	0.02	0.68	0.48
Digit Span Backward	0.05	0.00	0.29	0.80
Processing speed				
Trail Making Test: Part A	0.30	0.09	1.67	0.11
Digit-Symbol Coding	0.07	0.01	0.39	0.70
Verbal fluency				
Animals	0.32	0.10	1.79	0.08
Letter fluency	0.16	0.03	0.86	0.38

Note: For all analyses, $N = 30$.
"$r(x,y)$" Pearson correlation between the independent variable and the dependent variable; "r^2," coefficient of determination; "T" student's t-test; "p" probability.

viously, the diagnostic utility of this VRCPAT assessment tool must be determined. The ability of the VRCPAT to accurately classify participants into memory impaired and nonimpaired groups based on carefully established critical values must be evaluated. This will involve the generation of specific cut-off points for classifying a positive (memory impaired likely) or negative (memory impaired unlikely) finding. The VRCPAT's prediction of memory impairment must be evaluated by the performance indices of sensitivity, specificity, predictive value of a positive test, and predictive value of a negative test. Even though reliability is considered to be a unique asset of testing in computer-generated VEs, issues of test–retest reliability must be addressed.

Our goal was to conduct an initial pilot study to validate the VRCPAT through the use of a standard neuropsychological battery for the assessment of healthy participants. We believe that this goal was met. We recognize, however, that the current findings are only a first step in the development of this tool. Many more steps are necessary 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. Although the VRCPAT as a measure must be fully validated, current findings provide preliminary data regarding the validity of the VE 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 must be examined for cross-validation purposes.

REFERENCES

1. Tulving E. (2000) Concepts of memory. In: Tulving E, Craik FIM, eds., *The Oxford handbook of memory*. New York: Oxford University Press, pp. 33–44.
2. Squire LR. Mechanisms of memory. Science 1986; 232:1612–19.
3. Churchland PS, Sejnowski TJ. (1992) *The computational brain*. Cambridge, MA: MIT Press.
4. Kandel ER, Squire LR. Neuroscience: breaking down scientific barriers to the study of brain and mind. Science 2000; 290:1113–20.
5. Morris RGM, Moser EI, Riedel G, Martin SJ, Sandin J, Day M, et al. Elements of a neurobiological theory of the hippocampus: the role of activity-dependent synaptic plasticity in memory. Philosophical Transactions of the Royal Society of London—Series B: Biological Sciences 2003; 358:773–86.
6. Chaytor N, Schmitter-Edgecombe M, Burr R. Improving the ecological validity of executive functioning assessment. Archives of Clinical Neuropsychology 2006; 21:217–27.
7. Farias ST, Harrell E, Neumann C, Houtz A. The relationship between neuropsychological performance and daily functioning in individuals with Alzheimer's disease: ecological validity of neuropsychological tests. Archives of Clinical Neuropsychology 2003; 18:655–72.
8. Gioia GA, Isquith PK. Ecological assessment of executive function in traumatic brain injury. Developmental Neuropsychology 2004; 25:135–58.
9. Odhuba RA, van den Broek MD, Johns LC. Ecological validity of measures of executive functioning. The British Journal of Clinical Psychology 2005; 44:269–78.
10. Parsons TD, Rizzo AA, van der Zaag C, McGee JS, Buckwalter JG. Gender and cognitive performance: a test of the common cause hypothesis. Aging, Neuropsychology, & Cognition 2005; 12:78–88.
11. Parsons TD, Rizzo AA, Buckwalter JG. Backpropagation and regression: comparative utility for neuropsychologists. Journal of Clinical & Experimental Neuropsychology 2004; 26:95–104.
12. Parsons TD, Bowerly T, Buckwalter JG, Rizzo AA. A controlled clinical comparison of attention performance in children with ADHD in a virtual reality classroom compared to standard neuropsychological methods. Child Neuropsychology 2007.
13. Parsons TD, Larson P, Kratz K, Thiebaux M, Bluestein B, Buckwalter JG, et al. Sex differences in mental rotation and spatial rotation in a virtual environment. Neuropsychologia 2004; 42:555–62.
14. Wolbers T, Weiller C, Buchel C. Neural foundations of emerging route knowledge in complex spatial environments. Brain Research Cognitive Brain Research 2004; 21:401–11.
15. Matheis RJ, Schultheis MT, Tiersky LA, Deluca J, Millis SR, Rizzo A. Is learning and memory different in a virtual environment? Clinical Neuropsychology 2007; 21:146–61.
16. Elkind JS, 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 & Behavior 2001; 4:489–96.
17. Morganti F. Virtual interaction in cognitive neuropsychology. Studies in Health Technology & Informatics 2004; 99:55–70.
18. Gould NF, Holmes MK, Fantie BD, Luckenbaugh DA, Pine DS, Gould TD, et al. Performance on a virtual reality spatial memory navigation task in depressed patients. American Journal of Psychiatry 2007; 164:516–9.
19. Parslow DM, Morris RG, Fleminger S, Rahman Q, Abrahams S, Recce M. Allocentric spatial memory in

humans with hippocampal lesions. Acta Psycholog-ica 2005; 118:123–47.

20. Phelps A, Fritchle A, Hoffman H. Passive vs. active virtual reality learning: The effects on short- and long-term memory of anatomical structures. Studies in Health Technology & Informatics 2004; 98:298–300.

21. Attree EA, Brooks BM, Rose FD, Andrews TK, Lead-better AG, Clifford BR. (1996) Memory processes and virtual environments: I can't remember what was there, but I can remember how I got there. Implica-tions for people with disabilities. In: Sharkey P, ed., *Proceedings of the First European Conference on Disabil-ity, Virtual Reality and Associated Technology*. Maiden-head, UK: University of Reading, pp. 117–21.

22. Rose FD. (1996). Virtual reality in rehabilitation fol-lowing traumatic brain injury. In: Sharkey P, ed., *Pro-ceedings of the First European Conference on Disability, Virtual Reality and Associated Technology*. Maidenhead, UK: University of Reading, pp. 5–12.

23. Rose FD, Attree EA, Brooks BM, Andrews TK. Learn-ing and memory in virtual environments—a role in neurorehabilitation? Questions (and occasional an-swers) from UEL. Presence: Teleoperators & Virtual Environments 2001; 10:345–58.

24. Pugnetti L, Mendozzi L, Attree E, Barbieri E, Brooks BM, Cazzullo CL, et al. Probing memory and execu-tive functions with virtual reality: Past and present studies. CyberPsychology & Behavior 1998; 1:151–62.

25. Peruch P, Vercher JL, Gauthier GM. Acquisition of spa-tial knowledge through visual exploration of simulated environments. Ecological Psychology 1995; 7:1–20.

26. Wilson BA, Alderman N, Burgess PW, Emslie H, Evans JJ. (1996) *BADS: Behavioural Assessment of the Dysexecutive Syndrome*. Bury St Edmunds, UK: Thames Valley Test Company.

27. Wilson PN, Foreman N, Gillet R, Stanton D. Active versus passive processing of spatial information in a computer-simulated environment. Ecological Psy-chology 1997; 9:207–22.

28. Andrews TK, Rose FD, Leadbetter AG, Attree EA, Painter J. (1995) The use of virtual reality in the as-sessment of cognitive ability. In: Placencia Porrero I, Puig de la Bellacasa R, eds., *The European Context for Assistive Technology: Proceedings of the 2nd TIDE Con-gress*. Amsterdam: IOS Press.

29. Campbell DT, Fiske DW. Convergent and discrimi-nant validation by the multitrait-multimethod matrix. Psychological Bulletin 1959; 56:81–105.

30. Miller LS, Rohling ML. A statistical interpretive method for neuropsychological test data. Neuropsy-chology Review 2001; 11:141–67.

31. Kennedy RS, Fowlkes JE, Berbaum KS, Lilienthal MG. Use of a motion sickness history questionnaire for prediction of simulator sickness. Aviation, Space, & Environmental Medicine 1992; 63:588–93.

32. Brandt J, Benedict RHB. (2001) *Hopkins Verbal Learn-ing Test—revised. Professional manual*. Lutz, FL: Psy-chological Assessment Resources.

33. Benedict RHB. (1997). *Brief Visuospatial Memory Test—revised*. Odessa, FL: Psychological Assessment Re-sources.

34. Benton LA, Hamsher K, Sivan AB. (1994) Controlled Oral Word Association Test. In *Multilingual aphasia ex-amination* (3rd ed.). Iowa City, IA: AJA.

35. Gladsjo JA, Schuman CC, Evans JD, Peavy GM, Miller SW, Heaton RK. Norms for letter and category flu-ency: Demographic corrections for age, education, and ethnicity. Assessment 1999; 6:147–78.

36. Wechsler D. (2001) *Wechsler Test of Adult Reading (WTAR)*. New York: Psychological Corporation.

37. Heaton RK, Grant I, Matthews CG. (1991) *Compre-hensive norms for an expanded Halstead-Reitan battery: demographic corrections, research findings, and clinical ap-plications*. Odessa, FL: Psychological Assessment Re-sources.

38. Reitan RM, Wolfson D. (1985) *The Halstead-Reitan Neuropsychological test battery: THEORY and clinical in-terpretation*. Tucson, AZ: Neuropsychology Press.

39. Golden CJ. (1978) *Stroop Color and Word Test*. Chicago: Stoelting.

Address reprint requests to:
Dr. Thomas D. Parsons
University of Southern California
Institute for Creative Technologies
13274 Fiji Way, Office 301
Marina del Rey, CA 90292-4019

E-mail: tparsons@usc.edu