NeuroRehabilitation 41 (2017) 721–734
DOI:10.3233/NRE-172183
IOS Press

Application of virtual environments in a multi-disciplinary day neurorehabilitation program to improve executive functioning using the Stroop task

Marie N. Dahdah[a,b,c,*], Monica Bennett[d], Purvi Prajapati[d], Thomas D. Parsons[e], Erin Sullivan[f] and Simon Driver[a,b]
[a]*Baylor Institute for Rehabilitation, Dallas, TX, USA*
[b]*North Texas Traumatic Brain Injury Model System, Dallas, TX, USA*
[c]*Baylor Scott & White Medical Center – Plano, Plano, TX, USA*
[d]*Center for Clinical Effectiveness, Baylor Scott and White Health, Dallas, TX, USA*
[e]*Department of Psychology, Computational Neuropsychology and Simulation (CNS) Lab, University of North Texas, Denton, TX, USA*
[f]*Department of Psychology, University of North Texas, Denton, TX, USA*

Abstract.
BACKGROUND: Virtual reality (VR) technology has demonstrated usefulness in diagnosis, education, and training. Studies supporting use of VR as a therapeutic treatment in medical rehabilitation settings remain limited. This study examines the use of VR in a treatment capacity, and whether it can be effectively integrated into neurorehabilitation.
OBJECTIVE: To determine whether immersive VR treatment interventions improve executive dysfunction in patients with brain injury and whether performance is stronger on a VR version of the Stroop than traditional Stroop formats.
METHODS: 15 patients with brain injury admitted to day neurorehabilitation. Outcome measures: reaction time, inhibition, and accuracy indices on VR Stroop; Automated Neuropsychological Assessment Metrics (ANAM) Stroop, Delis-Kaplan Executive Function System Stroop, Golden Stroop, and Woodcock-Johnson, 3rd Edition (WJ-III): Pair Cancellation.
RESULTS: Participants demonstrated significantly reduced response time on the word-reading condition of VR Stroop and non-significantly reduced response time on the interference condition. Non-significant improvements in accuracy and inhibition were demonstrated on the color-naming condition of VR Stroop. Significantly improved accuracy under time pressure was found for the ANAM, after VR intervention.
CONCLUSION: Implementation of immersive VR interventions during neurorehabilitation is effective in improving specific executive functions and information processing speed in brain-injured patients during the subacute period.

Keywords: Brain injuries, virtual reality exposure therapy, executive function, treatment outcome, inpatients, rehabilitation

*Address for correspondence: Marie Dahdah, Ph.D, Neuropsychologist, Baylor Scott and White Health Medical Center - Plano, 4708 Alliance Blvd., Pavilion I, Suite 700, Plano, TX 75093, USA. Tel.: +1 469 814 4850; Fax: +1 469 814 5678; E-mail: Marie.Dahdah@BSWHealth.org.

1. Introduction

Initial applications of virtual reality (VR) in specific medicine and psychology specialties have demonstrated promise. Yet, little is known about how to effectively integrate VR into the rehabilitation realm as a means of improving care (Salisbury, Dahdah, Driver, Parsons, & Richter, 2016). VR permits experiential, active learning and multiple trials with the ability to (1) gradually increase the complexity of tasks, (2) maintain strict experimental control over stimulus delivery and measurement, and (3) individualize treatment needs in a standardized manner (Rand, Weiss, & Katz, 2009). VR also enables objective measurement of behaviors and functional outcomes in challenging but safe (ecologically valid) environments. Enhanced motivation has been demonstrated in patients with VR use (Larson, Feigon, Gagliardo, & Dvorkin, 2014; Rizzo & Kim, 2005) and a few studies utilizing a desktop display system with head mounted display (HMD) have demonstrated clinical effectiveness in cognitively impaired individuals (Hofmann et al., 2003; Simone, Schultheis, Rebimbas, & Millis, 2006).

Executive functions have traditionally been assessed using standardized neuropsychological measures that quantify neurobehavioral deficits (Cahn-Weiner, Boyle, & Malloy, 2002; Chaytor, Schmitter-Edgecombe, & Burr, 2006; Shallice & Burgess, 1991). However, conventional executive functioning tests require relatively simple responses to single events, whereas daily life tasks are complex and composed of multiple steps, such as formulation of goals and subgoals, prioritizing of subgoals, triggering prospective memory to initiate a subtask when the time is right, and inhibition of irrelevant and inappropriate actions to different subtasks (Chan, Shum, Toulopoulou, & Chen, 2008).

Introduction of VR technology into neuropsychological assessment and treatment has the potential to capitalize on the psychometric precision of traditional neuropsychological measures while utilizing the well-controlled properties of technologically advanced computerized measures (T. Parsons, 2011). VR environments allow for control of dynamic stimuli across different sensory modalities. In addition, the advanced computer interface permits accurate recordings of discrete neurobehavioral responses in a perceptual environment that closely approximates real-life scenarios and systematically presents complex stimuli (T. Parsons, 2011). Current virtual environments that have been studied in clinical populations include kitchens, school classrooms, university departments, cities, a street crossing environment, and a supermarket (Rand et al., 2009).

1.1. Virtual reality in assessment and diagnostic discrimination

Most clinical studies utilizing VR technology have used it for the purposes of assessment or diagnostic classification. Multiple studies have utilized a shopping task or grocery store simulation with clinical groups (Josman, Klinger, & Kizony, 2008; Kang et al., 2008). In one study examining executive functions in patients with mild cognitive impairment (MCI), stroke, and schizophrenia, participants were asked to purchase grocery items from a list and then pay for them at a cash register in the Virtual Action Planning-Supermarket (VAP-S) (Josman et al., 2008). Greater than 70% of the patients were correctly classified by diagnosis (Josman et al., 2008).

In another study, neurological patients completed a number of office-based tasks in a virtual office (Assessim Office), receiving manualized cueing if they became disoriented or lost attentional set for the tasks in the face of distracters (e.g. phone ringing) (Krch et al., 2013). Examples of tasks included printing real estate documents, making real estate decisions, or shutting off a projector light. For patients with traumatic brain injury (TBI), poor performance on the VR office tasks was significantly correlated with poor performance on a number of executive neuropsychological measures, while for MS patients the only significant correlation was with visual reasoning (Krch et al., 2013).

1.2. Virtual reality and rehabilitation

Fewer studies have used VR technology for treatment purposes, particularly in the rehabilitation milieu. In a study by Kim and his colleagues, patients with acute stroke all received computer-based cognitive rehabilitation comprised of programs designed to provide training in attention and memory functions, in addition to conventional physical and occupational therapies. However, some of these patients were also randomly assigned to a VR training condition, wherein they virtually re-positioned, smashed, punched, or blocked objects. Task complexity on the VR programs was adjusted based on patient deficits, and patients received performance feedback. Both groups demonstrated improvements from pre-treatment to post-treatment 4 weeks later. However,

stroke patients who received VR training demonstrated significant improvement in concentration and visual working memory compared to the control stroke group (Kim, Chun, Kim, & Park, 2011). In another study, patients with TBI were positioned on a recumbent bicycle they could maneuver in front of a VR screen (Grealy, Johnson, & Rushton, 1999). They received visual, auditory, tactile, and positioning cues, and performance feedback was provided. Neuropsychological testing was completed pre and post single-session intervention. Compared with archived data of age- and severity-matched patients with TBI who received multi-disciplinary rehabilitation, patients who received the VR exercise intervention demonstrated significantly stronger visual working memory, verbal and visual memory, and faster reaction time in comparison with controls (Grealy et al., 1999).

1.3. Executive performance on the Virtual Stroop in non-clinical samples

Studies comparing performance across different types of Stroop measures have been conducted in healthy individuals to-date. Twenty healthy college-aged participants were immersed in a virtual Iraqi/Afghani environment and were seated in a Humvee, with Stroop stimuli projected on the windshield of the vehicle (T. D. Parsons, Courtney, Arizmendi, & Dawson, 2011). Neurocognitive outcomes were evaluated in conjunction with physiological response (cardiac and electrodermal) to examine performance under different levels of threat (e.g. gunfire, explosions) (T. D. Parsons et al., 2011). A higher number of correct responses on the Stroop measure was achieved on a paper-and-pencil version and a computer-administered version (Automated Neuropsychological Assessment Metrics, ANAM) in comparison with the relatively more challenging VR Stroop task, the latter of which better measured inhibition in an immersive military environment. Also, the VR Stroop task discriminated between the more challenging interference condition versus the relatively simpler color-naming and word-reading conditions; the other Stroop versions did not (T. D. Parsons et al., 2011). This same VR Stroop task was validated on 49 active duty soldiers without posttraumatic stress disorder or TBI, aged 18–64 (Armstrong et al., 2013). Older patients responded more slowly on the Virtual Stroop word-reading and color-naming conditions than younger patients. The Virtual Stroop was found to be moderately correlated with a paper-and-pencil Stroop (Delis-Kaplan Executive Function System, D-KEFS: Color-Word Interference Test) and highly correlated with the ANAM. It took the soldiers significantly longer to complete the interference condition on all versions of the Stroop, compared with the word-reading and color-naming conditions (Armstrong et al., 2013).

VR technology has demonstrated its potential usefulness as a tool for diagnosis, education, and training. To date, application of VR in neurorehabilitation is not systematic and studies supporting the use of VR as a therapeutic treatment in medical rehabilitation settings remain limited. This study endeavors to address some of these limitations by using VR in a treatment capacity. Specifically, this study aims to examine: (1) whether systematic practice of a Stroop measure within an immersive VR environment improves executive functioning as environmental demands increase across treatment sessions, and (2) whether differences will be found between performance on paper-and-pencil, unimodal computerized, and bimodal VR versions of the Stroop task in a neurological population. Counter to findings in healthy individuals, it is expected that individuals with neurological insult will perform more poorly across all conditions of the Stroop when simulated environmental distraction is present.

2. Methods

2.1. Participants

Twenty-one brain injury (BI) patients with known executive dysfunction participated in this study: 9 diagnosed with stroke (43%), 6 with TBI (29%), 2 with anoxic injury (10%), 3 with brain tumor (14%), and 1 with amyloid angiopathy (5%). The final analyses included 15 participants, as 6 patients failed to complete all 8 intervention sessions. Specifically, 2 patients were medically withdrawn from the day neurorehabilitation program due to refractory medical complications, 2 patients self-discharged from the program against medical advice, and 4 patients' rehabilitation regimens were concluded prior to their projected discharge dates when insurance or state-assisted benefits were not extended.

2.2. Procedure

This study was conducted at a neurorehabilitation institution in the southern U.S. Patient participants

were those currently enrolled in an outpatient multidisciplinary outpatient neurorehabilitation program (Day Neuro). Data were obtained from consecutive admissions into the Day Neuro program from 08/2014 to 05/2015. The outpatient program is part of the continuum of care that includes a trauma center, inpatient rehabilitation facility, and comprehensive outpatient program. The program includes traditional rehabilitation care such as physical, occupational, recreation, and speech therapy as well as non-traditional services such as driving instruction, aquatics therapy, and home-based care (see Appendix 1 for greater detail). Individuals participate in Day Neuro once they have been discharged from the inpatient rehabilitation facility. Attendance is Monday through Friday, between 9am to 3pm. Patients participating in this study were aged 18 years and older and diagnosed with acquired traumatic or non-traumatic neurologic illness. The following exclusion criteria were applied: medically unstable, as deemed by the patient's primary doctor, absence of executive dysfunction (obtained from inpatient records and initial speech-language pathology evaluation) had prior history of significant neurological complications or developmental delay resulting in compromised cognition, prisoners, and non-English speakers. If a patient had not yet undergone a neuropsychological evaluation by the time they had consented to participate in the study, those patients were administered the Orientation and Cognitive Log (OLOG/Cog-Log) to ensure they were oriented and had sufficient cognitive ability to attend to and understand instructions for measures used in this study. Twenty-one patients met the inclusion criteria.

Patients completed the VR intervention sessions twice per week for a 4-week period (8 total sessions). Sessions were structured in the following way: Session 1 (baseline) included all types of distracters (auditory, visual, audio–visual) simultaneously. Sessions 2 and 3 included no distracting stimuli. To gauge whether the presence of distracters increased executive burden, distracters were then reintroduced at session 4, varying them by sensory modality. Specifically, sessions 4 and 5 included only auditory distracting stimuli, and sessions 6 and 7 included only visual distracting stimuli. Session 8 resembled baseline by including all types of distracters again, to gauge change in performance between sessions 1 and 8, after varying the presence and type of distracters across exposure trials. Appendix 1 provides greater detail regarding test conditions of the VR environments, distractions introduced, and all measures administered on each session day. The total duration of sessions 1 and 8 was approximately 60 minutes. The duration of sessions 2–7 was 30 minutes each. VR intervention days were planned on days separate from neuropsychological testing. Outside of the introduction of the 2 specified VR programs, clinical services were not altered.

Participants were fitted with a Z800 3D Visor head-mounted display system. The HMD was used to create a 3D-like effect allowing patients to look 360 degrees around themselves by turning their head. This HMD system has been approved by the U.S. Food & Drug Administration as part of a therapeutic application and has previously been used in rehabilitation patients with balance disorders, vertigo, or instability by the medical hardware developer Medicaa ("eMagin 3Dvisor™ Implemented by Medicaa Into FDA Approved Rehabilitation Unit," 2007). A laptop was placed central to the seated patient. Responses to the virtual reality and 2-dimensional computer measures were registered when the patient depressed the computer mouse.

2.3. Intervention description and measures

2.3.1. Bimodal VR-Stroop (ClinicaVR: Apartment Stroop)

Within this VR apartment environment, patients were seated in a living room, in front of a flat-screen TV set, a kitchen and a window (See Fig. 1). This intervention consisted of 2 Stroop conditions across all 8 sessions. In Condition 1 (Inhibition), a series of color rectangles appeared on the television screen (blue, red or green) while the name of one of these colors was verbally recited through the computer speakers by a female voice at the same pace (bimodal presentations). Participants were expected to click on the left button of a mouse with their preferred hand as quickly as possible when the color named (audio stimulus) matched the color shown (visual stimulus). They were instructed to withhold their response in mismatched trials. A total of 144 stimuli were presented, including 72 targets (Henry, Joyal, & Nolin, 2012). During the task, 14 distracters appeared in different areas of the environment (center, left, or right). Some distracters were *audio–visual* (School Bus passing on the street, Toy Robot on the floor), others were *auditory* (Doorbell, Vacuum Cleaner), and some were *visual* (Paper plane, Woman Walking in Kitchen). Distracters were displayed for 5 seconds,

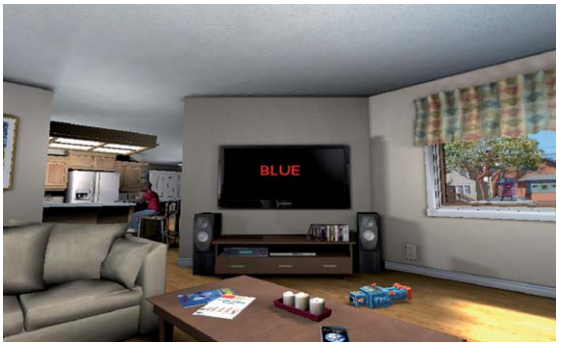

Fig. 1. Virtual Reality Apartment. Digital Media Works: Kenata, ON, Canada, www.dmw.ca.

and presented in equally appearing intervals of 10, 15, or 25 seconds. In Condition 2 (Interference), color words were presented on the screen, written with matched ink color (e.g. BLUE written in blue, congruent trial) or different ink color (e.g. BLUE written in red, incongruent trial). Participants were instructed to click on the mouse when the color heard was the same as the ink color; not the word printed. Again, a total of 144 stimuli were presented, including 72 targets, divided into 36 congruent and 36 incongruent stimuli. Distracters in Condition 2 were the same as those in Condition 1. Total task duration, including both conditions, was 9.6 minutes. Outcomes recorded included: (1) mean reaction times for correct and incongruent trials; (2) mean, shortest, and longest response times for correct responses; (3) mean, shortest, and longest response time for incongruent stimuli; (4) total commission errors; and (5) total omission errors.

2.3.2. *Bimodal VR-Stroop (VR Classroom)*

The Virtual Classroom was selected given that some rehabilitation patients report a goal of returning to school, and because elements of the environment also mimic an office setting. It was hoped that this would make the Virtual Classroom dually useful for patients hoping to resume specific types of office work post-discharge.

This environment presented the interior of a standard rectangular classroom environment containing 3 rows of desks in the HMD. A teacher's desk was located at the front of the classroom, a blackboard across the front wall, and a female virtual teacher was standing between the desk and blackboard. On the left side wall, there was a large window looking out onto a playground with buildings, vehicles, and people. On each end of the wall opposite the window, there was a pair of doorways through which activity occurred. The task conditions, participant requirements, and task duration were the same as those in the VR apartment. The only difference was the distracters: *audio-visual* distracters (car passing outside of the window; and a man entering and exiting creaking doors), *auditory* distracters (whispering, chairs moving), and *visual* distracters (3D paper airplane flying in front of the participant) (T. D. Parsons, Bowerly, Buckwalter, & Rizzo, 2007). The same outcomes were recorded as for VR apartment.

2.4. *Outcome measures*

To evaluate improvement in executive functions the Woodcock-Johnson, 3rd Edition (WJ-III): Pair Cancellation subtest was used to measure visual selective attention, sustained attention and concentration, as well as visuomotor processing speed. The Delis-Kaplan Executive Function System (D-KEFS): Color-Word Interference subtest and Automated Neuropsychological Assessment Metrics (ANAM): Go/No-Go and unimodal Stroop subtests evaluated processing speed, inhibition, and cognitive flexibility. These 3 measures were administered during baseline (Session 1) and Session 8, when all distractors were included.

2.5. *Symptom Self-Report Questionnaire*

The Simulator Sickness Questionnaire (SSQ) was completed by patients upon conclusion of Sessions 2 through 7 to assess the occurrence, nature, and severity of sickness symptoms induced by VR environments. The SSQ comprises 16 items rated on a scale from 0 to 3 (values are based on how much a particular symptom is affecting a patient with 0 = none; 1 = slight; 2 = moderate; 3 = severe).

2.6. *Statistical analysis*

All analyses were conducted using SAS, version 9.4 (SAS Institute Inc, Cary, North Carolina). Two-tailed p-values ≤ 0.05 were considered statistically significant. Descriptive statistics were summarized as means and standard deviations or medians and interquartile ranges for continuous variables, and percentages and frequencies for categorical variables. Paired t-tests were employed to make comparisons between VR Session 1 and Session 8 variables, as well as Session 1 and 2 variables. The relationships between different versions of Stroop and VR

Table 1
Demographic and clinical information of Virtual Reality participants

	Virtual Reality ($n = 15$)
Age, mean (sd)	40.3 (16.1)
Gender	
Male	12 (80%)
Female	3 (20%)
Ethnicity/Race	
Black	3 (20%)
Hispanic	3 (20%)
White	9 (60%)
Years of Education	
<12	2 (13.3%)
12–16	12 (80.0%)
16+	1 (6.7%)
Diagnosis	
CVA	6 (40%)
Traumatic brain injury	5 (33%)
Tumor	2 (13%)
Anoxia brain injury	2 (13%)
Time Variables	
Medical LOS (days), median (IQR)	13.5 (9–22)
Rehab LOS (days), median (IQR)	26.5 (15–34)
Day Neuro (days), median (IQR)	51 (28–83)

Note: LOS = length of stay; IQR = interquartile range.

variables were summarized using bivariate Pearson correlations, and were examined to determine convergent and discriminant validity (Cohen, 1992). All versions of the Stroop task (VR, ANAM, D-KEFS, Golden) were converted to milliseconds to make them equivalent across formats. Based upon prior research, multiple patient demographics and clinical factors that may influence functional outcomes were included in analyses conducted in this study. These included age, sex, race/ethnicity, marital status, years of education, and length of stay (Brown et al., 2005; Cowen et al., 1995; Frankel et al., 2006; Ratcliff et al., 2007; Schopp, Shigaki, Johnstone, & Kirkpatrick, 2001).

3. Results

Table 1 summarizes patient demographics and length of medical, inpatient rehabilitation, and day neurorehabilitation stay (LOS). Mean age of the study population was 40 years. Eighty percent of the sample were men, 40% were non-White race/ethnicity, and most participants had completed at least a high school education (86%). The majority of participants were diagnosed with stroke (40%) or TBI (33%) and median LOS was nearly 2 months in the Day Neuro program.

3.1. Performance outcomes on different versions of the Stroop

Response times were recorded in milliseconds across all versions of the Stroop. No significant change in performance was found for the paper-pencil version of the Stroop (D-KEFS) when comparing baseline performance with performance following 8 sessions of VR intervention, with the exception of a non-significant trend towards less time taken to complete the color-naming condition by session 8 (Session 1 = 698 ms; Session 8 = 634 ms; $p = 0.08$). Participants achieved significantly higher accuracy on the word-reading trial of the unimodal ANAM Stroop after 8 sessions of VR intervention compared with baseline performance, when time constraints were factored in (see Table 2). Participants committed significantly fewer commission errors on the ANAM Go/No-Go subtest by session 8 of the study (see Table 2).

3.2. VR performance at conclusion of treatment

3.2.1. Conditions with distractors only

No statistically significant performance differences were found from baseline to conclusion of the study for the VR apartment variables. However, a non-significant trend towards improved accuracy (total number of correct responses: Session 1 = 70 ms; Session 8 = 71 ms; $p = 0.08$) and reduction in omission errors (Session 1 = 1.1 errors; Session 8 = 0.6 errors; $p = 0.08$) was observed for the color naming condition. For the VR classroom, participants' shortest response time on the word-reading condition was significantly reduced by Session 8 (see Table 3). There was a non-significant trend towards abbreviation of this same variable over time for the interference condition (Session 1 = 562 ms; Session 8 = 453 ms; $p = 0.09$).

3.2.2. Comparison with and without distractors

For the VR apartment, participants' reaction time was significantly faster when no distracters were present during the color-naming condition, specifically when longer response times for correct responses were assessed (see Table 4). Average response time associated with commission errors was longer (though not significantly, $p = 0.09$) in the presence of distracters for the interference condition.

For the VR classroom, participants spent significantly less time considering their response on incongruent items of the interference condition when

Table 2
Automated Neuropsychological Assessment Metrics (ANAM) Stroop and Go/No-Go performance in Session 1 and Session 8 for patients receiving Virtual Reality treatment (N = 15)

Variable Description	Session 1	Session 8	*p*-value
Stroop Word-Reading			
Total number of correct responses (including correct non-responses)	28.7 ± 7.5	31.2 ± 7.5	0.0741
Percentage of items with a correct response	61.5 ± 24	70.3 ± 25.4	**0.0293**
Number of correct responses per minute	61.5 ± 24	70.1 ± 25.2	**0.0321**
Stroop Color-Naming			
Total number of correct responses (including correct non-responses)	29.6 ± 9.9	31.4 ± 7.7	0.4226
Percentage of items with a correct response	73.5 ± 22.1	74.8 ± 26.2	0.7913
Number of correct responses per minute	73.5 ± 22.1	74.8 ± 26.2	0.7913
Stroop Interference			
Total number of correct responses (including correct non-responses)	24.6 ± 7.7	25.2 ± 10.3	0.6825
Percentage of items with a correct response	50.8 ± 21.3	54.6 ± 28.5	0.3499
Number of correct responses per minute	50.8 ± 21.3	54.6 ± 28.5	0.3499
Go/No-Go			
Average response time for items with correct responses	437.6 ± 85.3	462.4 ± 108.6	0.7100
Average response time of items with incorrect responses	232.9 ± 128.5	290.9 ± 134.9	0.448
Number of impulsive/bad responses (RT < 130 ms)	95.1 ± 2.9	92.3 ± 4.6	**0.0408**

Table 3
Change in performance between Session 1 and Session 8 on the Virtual Reality Classroom Stroop (auditory and visual distractors included in both sessions)

Variable Description	Session 1	Session 8	*p*-value
Color Naming			
Total number of correct responses	71.3 ± 1.6	69.5 ± 5.1	0.2132
Average response time for correct responses	741.2 ± 113.2	752.9 ± 154.8	0.6869
Shortest response time for correct responses	488.6 ± 97.2	530.5 ± 83.1	0.2478
Total number of omissions	0.7 ± 1.6	2.3 ± 5	0.2514
Word Reading			
Total number of correct responses	61.5 ± 13.6	61.8 ± 13.9	0.9482
Average response time for correct responses	816.8 ± 101.2	788.3 ± 125.4	0.2654
Shortest response time for correct responses	553.2 ± 75.8	502.2 ± 102.3	**0.0383**
Mean response time for congruent stimuli	766.3 ± 91	750.8 ± 135.6	0.5742
Interference			
Total number of correct responses	26.3 ± 13.1	27.9 ± 12.2	0.6536
Average response time for incongruent stimuli	776.6 ± 338	731.4 ± 320.8	0.626
Shortest response time for incongruent stimuli	562.3 ± 243.9	453.4 ± 212.3	0.0918
Total number of omissions	10.4 ± 13.6	10.1 ± 13.8	0.9477

distracters were absent (see Table 5). Contrary to findings for the VR apartment, response time was not significantly longer when commission errors were made in the presence of distracters for the interference condition.

No significant change in performance was found for the Pair Cancellation measure from baseline to session 8, with the exception of a non-significant trend towards improved selective visual attention, reflected in relatively improved accuracy in identifying target visual stimuli by session 8 (see Table 6).

Ten of the 16 items were positively endorsed by patients on the SSQ across sessions (see Table 7). General discomfort, fatigue, eye strain, difficulty focusing or concentrating, and blurred vision of mild or moderate degree were mostly consistently endorsed across sessions. By the fourth session, the number of patients reporting HMD-related symptoms declined to four. No adverse events occurred and no patients volitionally withdrew from the study.

3.2.3. *Convergent and discriminant validity: Comparison of different versions of the Stroop*

Average response time on the interference conditions of the VR apartment and classroom was faster than the average response time for the D-KEFS and ANAM version of the Stroop (see Table 8). Tukey's pairwise comparisons revealed that response time on the ANAM Stroop was significantly longer than on all other Stroop measures.

There was a significant association between total number of incorrect responses on the VR apartment color-naming condition and fewer correct responses

Table 4
Virtual Reality Apartment Stroop interference variables: Comparison of performance with and without distractors (N = 14)*

Virtual Reality Apartment Stroop Variable	Session 1 All Distractors	Session 2 No Distractors	*p*-value
Color Naming			
Total number of correct responses	70.8 ± 1.3	70.6 ± 1.2	0.7202
Average response time for correct responses	760.3 ± 132.8	719.3 ± 144.7	0.2816
Shortest response time for correct responses	491.6 ± 163	474.2 ± 72.6	0.5913
Longest response time for correct responses	1380.5 ± 315.3	1170.5 ± 335.6	**0.036**
Average response time for incorrect responses	432.8 ± 371.8	350.3 ± 452.2	0.4988
Total number of omissions	1.1 ± 1.1	1.3 ± 1	0.6714
Word Reading			
Total number of correct responses	61.8 ± 12.2	63.9 ± 12.2	0.4894
Average response time for correct responses	830 ± 137.7	820 ± 148	0.7763
Shortest response time for correct responses	500.6 ± 146	523.1 ± 78.3	0.622
Longest response time for correct responses	1402.4 ± 301.2	1570.2 ± 220.3	0.0958
Average response time for congruent stimuli	786.3 ± 148.4	774.6 ± 142.7	0.7389
Total number of correct responses to incongruent stimuli	61.8 ± 12.2	63.9 ± 12.2	0.4894
Interference			
Total number of correct responses	27 ± 12	28.9 ± 12.5	0.5367
Average response time for correct responses	745.5 ± 335.3	755.1 ± 355.5	0.9232
Shortest response time for correct responses	496.1 ± 227.2	524.2 ± 262.2	0.6749
Longest response time for correct responses	1215.5 ± 586.8	1384.2 ± 610.6	0.3265
Average response time for incorrect responses	780.9 ± 400.2	578 ± 404.6	0.0983
Total number of omissions	69.1 ± 2.3	69.2 ± 2.8	0.8555

*One participant was unable to complete Session 2. The N for this analysis is 14 participants.

Table 5
Virtual Reality Classroom Stroop interference variables: Comparison of performance with and without distractors (N = 14)*

Virtual Reality Classroom Stroop Variable	Session 1 All Distractors	Session 2 No Distractors	*p*-value
Color Naming			
Total number of correct responses	71.4 ± 1.6	69.6 ± 3.7	0.1092
Average response time for correct responses	746.8 ± 115.3	754.4 ± 179.5	0.8495
Shortest response time for correct responses	486.7 ± 100.5	524.3 ± 124.3	0.38
Longest response time for correct responses	1354.8 ± 344.5	1303.6 ± 369.6	0.5268
Average response time for incorrect responses	359.6 ± 362.1	317.6 ± 339.2	0.758
Total number of omissions	0.6 ± 1.6	1.8 ± 2.6	0.1684
Word Reading			
Total number of correct responses	63.5 ± 11.6	63.8 ± 11.3	0.9186
Average response time for correct responses	820.5 ± 104	863.9 ± 198.5	0.2637
Shortest response time for correct responses	550.2 ± 77.7	486.6 ± 171.2	0.1082
Longest response time for correct responses	1471.3 ± 266.8	1533.3 ± 242.7	0.383
Total number of correct responses for congruent stimuli	35.3 ± 1.1	33.4 ± 4.4	0.1607
Interference			
Total number of correct responses	28.2 ± 11.3	30.4 ± 9.5	0.2989
Average response time for correct responses	832.1 ± 270.9	859.6 ± 312.9	0.4097
Shortest response time for correct responses	602.4 ± 194.9	468.5 ± 233	**0.0229**
Longest response time for correct responses	1341.9 ± 477.4	1370.7 ± 480.5	0.7015
Average response time for incorrect responses	621.7 ± 434.6	767.8 ± 507.6	0.2215
Total number of omissions	8.4 ± 11.7	8.1 ± 11.2	0.918

*One participant was unable to complete session 2. The N for this analysis is 14 participants.

Table 6
Pair Cancellation performance in Session 1 and Session 8 for patients receiving Virtual Reality treatment (N = 15)

Variable Description	Session 1	Session 8	*p*-value
Pair Cancellation Number Correct	53.1 ± 12.3	57.5 ± 13	0.0697
Pair Cancellation Standard Score	85.3 ± 15	88.1 ± 16.1	0.1226
Pair Cancellation Time in seconds	162.7 ± 20.7	165.8 ± 22.1	0.6021

Table 7
Simulator Sickness Questionnaire (SSQ) endorsements during Sessions 2 through 7 (N = 15)

	Virtual Reality Patients *n* (%)
Simulator Sickness Questionnaire Item #	
SSQ 1 – General Discomfort	3 (20%)
SSQ 2 – Fatigue	6 (40%)
SSQ 3 – Headache	1 (6.7%)
SSQ 4 – Eyestrain	4 (26.7%)
SSQ 5 – Difficulty Focusing	4 (26.7%)
SSQ 6 – Salivation Increasing	0 (0%)
SSQ 7 – Sweating	1 (6.7%)
SSQ 8 – Nausea	0 (0%)
SSQ 9 – Difficulty Concentrating	4 (26.7%)
SSQ 10 – Fullness of the Head	0 (0%)
SSQ 11 – Blurred Vision	3 (20%)
SSQ 12 – Dizziness with Eyes Open	1 (6.7%)
SSQ 13 – Dizziness with Eyes Closed	0 (0%)
SSQ 14 – Vertigo	1 (6.7%)
SSQ 15 – Stomach Awareness	0 (0%)
SSQ 16 – Burping	0 (0%)

achieved within standardized time limits on the paper-pencil Golden Stroop color-naming condition (see Table 9). A higher number of omission errors on the color-naming condition of the VR apartment corresponded with faster response time on the Golden Stroop color-naming condition. Faster mean response times on the word-reading condition of the VR apartment were associated with a greater number of correct responses achieved within standardized time constraints on the paper-pencil D-KEFS Stroop. However, as response times became increasingly shorter (faster), this was associated with poorer performance on the D-KEFS Stroop. On the interference condition of the VR apartment, longer response times associated with commission errors (color heard did not match color of word on flat screen TV in VR environment) was significantly correlated with poorer performance on the paper-pencil D-KEFS Stroop interference trial (color response by patient did not match color of printed word), and longer response times in general on the D-KEFS.

There were no significant associations between the bimodal VR apartment Stroop and the unimodal ANAM Stroop.

4. Discussion

Individuals with brain injury in this study demonstrated improved sustained attention, attention to visual details, cognitive flexibility, and relatively fewer impulsive errors across sessions of a VR intervention that resembled one's home environment, as reflected in improved accuracy and fewer commission errors. Information processing speed (faster reaction time) improved across sessions of a VR intervention that resembled a school/work environment. Interestingly, average response time for the interference condition of the VR measures was faster relative to response times on the D-KEFS and ANAM Stroop versions, and even relative to the less cognitively taxing VR word-reading conditions. This is counter to findings in healthy active duty soldiers, who demonstrated longer completion times on the interference condition of the same three versions of the Stroop administered in this study, compared with word-reading and color-naming conditions (Armstrong, Reger et al., 2013; Parsons, Courtney, et al., 2011). This is likely due to the relatively more complex attentional demands of the interference condition and the presence of impulsivity and reduced set-shifting for some neurologic patients with executive dysfunction.

In spite of these trends, no statistically significant differences were found for the VR apartment variables. For the VR classroom, only word-reading response time was significantly reduced over time, which is considered relatively easiest among the 3 Stroop conditions. Additionally, response time was faster when giving incorrect responses, which is contrary to findings relating to cognitive conflict resolution in healthy individuals (Nombela et al., 2014). This is likely due to dysexecutive impulsivity in this neurological population, where errors are likely due to inadequate time spent deliberating on error items. The significant findings for VR classroom only may suggest a possible double-practice effect. Participants completed the VR classroom intervention 2nd during each treatment session. It is possible that the

Table 8
Comparisons between the average response times for D-KEFS Stroop, Golden Stroop, and Automated Neuropsychological Assessment Metrics (ANAM) Stroop vs. Virtual Reality (VR) Apartment Stroop using analysis of variance (ANOVA) and Tukey's pairwise comparisons

	Stroop Measure				
Stroop Condition	VR Apartment	VR Classroom	ANAM Stroop	D-KEFS Stroop	*p*-value
Color	752.8 ± 131.2	741.2 ± 113.1	1052.8 ± 406.8	698.7 ± 199.9	0.0007
Word	825.2 ± 134	816.8 ± 101.2	881.1 ± 317.9	526.7 ± 125.1	<0.0001
Interference	816.8 ± 101.2	776.6 ± 338	1327.7 ± 684.9	1190.7 ± 321.2	0.001

Table 9
Bivariate correlations between D-KEFS Stroop, Golden Stroop, and Automated Neuropsychological Assessment Metrics (ANAM) Stroop vs. Virtual Reality (VR) Apartment Stroop

	DKEFS Stroop Mean Time Per Response (r value)	Golden Stroop Mean Time Per Response (r value)	ANAM Stroop	
			Mean Response Time for Correct Items (r value)	Mean Response Time for Incorrect Items (r value)
Color Naming				
Total number of correct responses	0.011	–0.078	0.187	–0.049
Average response time for correct responses	0.560	–0.165	0.538	–0.509
Shortest response time for correct responses	0.151	–0.144	0.532	–0.289
Total number of incorrect responses	0.425	**0.605***	0.083	–0.060
Average response time for incorrect responses	0.334	0.484	0.072	0.026
Longest response time for an incorrect response	0.395	0.554	0.029	–0.011
Stimuli that were correct unanswered	–0.425	**–0.605***	–0.083	0.060
Total number of omissions	–0.039	0.075	–0.152	0.017
Word Reading				
Total number of correct responses	0.142	0.304	–0.010	0.221
Average response time for correct responses	**0.735**†	0.062	0.408	–0.297
Shortest response time for correct responses	–0.562	0.385	–0.565	0.088
Total number of correct responses for congruent stimuli	0.059	–0.543	0.259	–0.004
Mean response time for congruent stimuli	**0.670***	–0.012	–0.559	–0.175
Shortest response time for congruent stimuli	**–0.606***	0.386	0.009	0.063
Longest response time for congruent stimuli	0.319	0.343	0.025	–0.073
Interference				
Total number of correct responses	0.257	0.310	–0.174	–0.184
Average response time for correct responses	0.131	0.441	–0.211	–0.314
Shortest response time for correct responses	–0.073	0.334	–0.397	–0.409
Total number of incorrect responses	0.200	0.142	0.042	–0.007
Average response time for incorrect responses	**0.634***	0.136	0.426	0.114
Shortest response time for an incorrect response	0.365	–0.094	0.315	0.048
Longest response time for an incorrect response	**0.638***	0.281	0.341	0.069
Stimuli that were correct unanswered	–0.200	–0.142	–0.042	0.007
Total number of omissions	–0.202	–0.247	0.247	0.188

*Significant at 0.05. †Significant at 0.01.

opportunity to practice these neurocognitive functions multiple times within and across sessions is what contributed to statistically detectable benefit over the course of rehabilitation.

Intuitively, response times on the VR interventions were longer in the presence of distracters for specific indices of all Stroop conditions (word-reading, color-naming, interference). Interestingly, average response time on the interference conditions of the VR apartment and classroom was faster than the average response time for the D-KEFS and ANAM version of the Stroop. The reverse was found in healthy individuals, whose mean response time was significantly longer on VR Stroop compared with ANAM and D-KEFS versions of the Stroop (Armstrong et al., 2013). This is an important finding, and it supports the ecological validity of immersive VR environments in capturing neurocognitive deficits in patients with neurological dysfunction. Patients were more likely to respond impulsively or deliberate less on their responses when in an environment mimicking their home or work, and when distractors were impinging on their attentional reserve. In contrast, other versions of the Stroop administered in this study are relatively more structured. There is no change in background features on ANAM, D-KEFS, or Golden versions of the Stroop, and no distractors are incorporated into their standardized administration. This finding is further supported by the fact that commission errors and faster reaction time on VR interventions was associated with poorer performance on paper-pencil versions of the Stroop (D-KEFS and Golden), considered to be relatively least cognitively taxing of all versions administered to patients in this study.

Studies have shown that older individuals exhibit greater sensitivity to visual distraction in multiple-item testing procedures, as compared with a single item being presented at a time, impacting performance (Davidson, Zacks, & Williams, 2003; Lustig, Tonev, & Hasher, 2000). Similar findings were observed in this study of neurological

patients, in that statistically detectable change in performance was observed for the ANAM measures (Stroop and Go/No-Go), which present a single stimulus on a plain black screen, but not consistently for the measures with high visual stimulus volume (VR, D-KEFS Stroop, WJ-III Pair Cancellation).

Findings from this study demonstrated that VR environments could be easily incorporated into the neurorehabilitation milieu. It also provided preliminary evidence for the use of VR in treating deficits in executive functions and processing speed, even though not all findings reached significance. With a greater sample size to improve power and perhaps an increased number of sessions, it is possible that stronger statistical differences may have been realized by conclusion of the study. Follow-up study is necessary to evaluate these factors. Improvements across time in this study were not attributable to spontaneous recovery, given that patients were at least 3 months post-neurologic insult. A limitation noted in multiple VR studies is the absence of a control group (Larson et al., 2014). Analyses comparing differences in neurorehabilitation outcomes between the group in this study and a demographically matched group of patients with brain injury is currently underway by this group.

The desktop display system and HMD used to create these immersive environments also has functional benefit, by reducing treatments barriers present for older individuals, patients with significant physical limitations, or patients without transportation (Cherniak, 2011). With an increasing number of studies demonstrating benefit of VR in improving neurocognitive functions in a dynamic fast-paced rehabilitation or hospital setting, this may open the door to specialized neurocognitive rehabilitation treatments being offered in a home or other institutional setting. This has further positive implications for potentially reduced hospital costs and increased generalizability of rehabilitation gains.

4.1. Study limitations

There are several limitations to consider in this study. When this study was first conceived, it was expected that most patients in the Day Neuro program would have diagnoses of TBI and stroke based on historical enrollment trends, and that a sufficient number of patients would be enrolled to permit examination of differences in VR performance for these 2 groups. Clinical factors outside of the control of these researchers (e.g. abrupt illness, loss of funding for rehabilitation) and exclusion of individuals with insufficient cognitive capacity to participate in this study contributed to the low sample size. However, it should be noted that most VR studies are composed of fewer than 30 total participants. Due to the small sample size, low power may have impacted our ability to detect significant differences in performance from Session 1 to 8. This also prevented us from examining differences by etiology. The groups consisted of patients with various neurological etiologies and severities of neurologic insult. Most studies administer various embellishments of the VR Stroop during a single administration. No other study has used it as a training intervention to-date. As such, it is yet unclear whether additional sessions on the VR interventions in this study would have demonstrated even greater improvement in executive functions and processing speed. This should be considered in the context of time-limited outpatient rehabilitation courses.

The use of a desktop display system with HMD may limit the availability, cost, time for set up, and time and resources necessary for training clinicians to use such equipment regularly during rehabilitation. Although use of this equipment was not a problem for participants in this study, this was a relatively younger group. It is possible that older individuals without computer experience may have been challenged by the use of this newer technology. Though few participants reported symptoms associated with HMD use, it was possible that due to cognitive deficits their reports were confounded with their general condition associated with CNS injury, as opposed to reporting new onset symptoms associated with simulator sickness. It was noted that significant performance differences were observed more for VR classroom indices than for the VR apartment, despite use of precisely the same executive tasks and administration procedures. The protocol intended to counterbalance order of VR apartment and classroom administrations. However, given that data collection was completed before this took place, there was a possibility of order effects, such that completion of the VR version of the Stroop task twice in one session may have resulted in a double practice-effect for VR classroom. Additional studies are needed to assess whether differences in the 2 VR environments and distractors in the respective environments may have contributed to these performance differences.

Rehabilitation treatment frequency/ intensity, participation, degree of psychosocial support of the

patient, the patient's emotional well-being, types of specialty services offered (e.g. educational groups), and preexisting illnesses were not measured or analyzed in this study. Thus, additional factors may explain findings from this study. Findings from this group of patients may not necessarily generalize to other neurorehabilitation populations.

5. Conclusion

Use of immersive VR interventions that simulate natural settings were incorporated into a neurorehabilitation regimen without significant difficulty, they were palatable to patients, and they resulted in no significant aversive symptoms. VR environments utilizing a Stroop task for serial treatment were effective in improving specific executive functions of complex attention, cognitive flexibility, inhibition, and information processing speed in brain-injured patients, which could not be purely attributed to spontaneous recovery during the subacute period. Further research is necessary to establish specific treatment protocols incorporating such VR interventions, in order to provide guidance on frequency, intensity, and how to pair these interventions with other elements of conventional neurorehabilitation to improve the effectiveness and efficiency of neurorehabilitation.

Acknowledgments

The authors would like to thank the Baylor Institute for Rehabilitation's Day Neurorehabilitation program in Dallas for their coordination efforts throughout their patients' time in this study. We also express our gratitude to Digital Media Works for use of their VR programs.

Conflict of interest

The contents of this manuscript were developed under a grant from the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR grant number 90DP0045-01-0). NIDILRR is a Center within the Administration for Community Living (ACL), Department of Health and Human Services (HHS). The contents of this manuscript do not necessarily represent the policy of NIDILRR, ACL, HHS, and you should not assume endorsement by the Federal Government. The project was also supported by the Ginger Murchison Foundation. There is no financial conflict of interest to report, as no financial benefit will be conferred on the authors or on any organization with which the authors are associated as a result of the research supporting this article.

References

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. *J Clin Exp Neuropsychol*, *35*(2), 113-123. doi: 10.1080/13803395.2012.740002

Brown, A. W., Malec, J. F., McClelland, R. L., Diehl, N. N., Englander, J., & Cifu, D. X. (2005). Clinical elements that predict outcome after traumatic brain injury: A prospective multicenter recursive partitioning (decision-tree) analysis. *J Neurotrauma*, *22*(10), 1040-1051. doi: 10.1089/neu.2005.22.1040

Cahn-Weiner, D. A., Boyle, P. A., & Malloy, P. F. (2002). Tests of executive function predict instrumental activities of daily living in community-dwelling older individuals. *Appl Neuropsychol*, *9*(3), 187-191. doi: 10.1207/S15324826AN0903_8

Chan, R. C., Shum, D., Toulopoulou, T., & Chen, E. Y. (2008). Assessment of executive functions: Review of instruments and identification of critical issues. *Arch Clin Neuropsychol*, *23*(2), 201-216. doi: 10.1016/j.acn.2007.08.010

Chaytor, N., Schmitter-Edgecombe, M., & Burr, R. (2006). Improving the ecological validity of executive functioning assessment. *Arch Clin Neuropsychol*, *21*(3), 217-227. doi: 10.1016/j.acn.2005.12.002

Cherniak, E. P. (2011). Not just fun and games: Application of virtual reality in the identification and rehabilitation of cognitive disorders in the elderly. *Disability and Rehabilitation: Assistive Technology*, *6*(4), 283-289.

Cohen, J. (1992). A power primer. *Psychol Bull*, *112*(1), 155-159.

Cowen, T. D., Meythaler, J. M., DeVivo, M. J., Ivie, C. S., 3rd, Lebow, J., & Novack, T. A. (1995). Influence of early variables in traumatic brain injury on functional independence measure scores and rehabilitation length of stay and charges. *Arch Phys Med Rehabil*, *76*(9), 797-803.

Davidson, D. J., Zacks, R. T., & Williams, C. C. (2003). Stroop interference, practice, and aging. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn*, *10*(2), 85-98. doi: 10.1076/anec.10.2.85.14463

eMagin 3DVisor(TM) Implemented by Medicaa Into FDA Approved Rehabilitation Unit. (2007). Retrieved February 16, 2015, from http://www.businesswire.com/news/home/20070321005290/en/eMagin-3DVisor-TM-Implemented-Medicaa-FDA-Approved-Rehabilitation#.U2mk2U1eGB1

Frankel, J. E., Marwitz, J. H., Cifu, D. X., Kreutzer, J. S., Englander, J., & Rosenthal, M. (2006). A follow-up study of older adults with traumatic brain injury: Taking into account decreasing length of stay. *Arch Phys Med Rehabil*, *87*(1), 57-62. doi: 10.1016/j.apmr.2005.07.309

Grealy, M. A., Johnson, D. A., & Rushton, S. K. (1999). Improving cognitive function after brain injury: The use of exercise and virtual reality. *Arch Phys Med Rehabil*, *80*(6), 661-667.

Henry, M., Joyal, C. C., & Nolin, P. (2012). Development and initial assessment of a new paradigm for assessing cognitive and motor inhibition: The bimodal virtual-reality Stroop. *J Neurosci Methods*, *210*(2), 125-131. doi: 10.1016/j.jneumeth.2012.07.025

Hofmann, M., Rosler, A., Schwarz, W., Muller-Spahn, F., Krauchi, K., Hock, C., & Seifritz, E. (2003). Interactive computer-training as a therapeutic tool in Alzheimer's disease. *Compr Psychiatry*, *44*(3), 213-219. doi: 10.1016/S0010-440X(03)00006-3

Josman, N., Klinger, E., & Kizony, R. (2008, September 10-12, 2012). *Performance within the Virtual Action Planning Supermarket (VAP–S): An executive function profile of three different populations suffering from deficits in the central nervous system.* Paper presented at the Proceedings of the 7th International Conference on Disability, Virtual Reality and Associated Technologies, Laval, France.

Kang, Y. J., Ku, J., Han, K., Kim, S. I., Yu, T. W., Lee, J. H., & Park, C. I. (2008). Development and clinical trial of virtual reality-based cognitive assessment in people with stroke: Preliminary study. *Cyberpsychol Behav*, *11*(3), 329-339. doi: 10.1089/cpb.2007.0116

Kim, B. R., Chun, M. H., Kim, L. S., & Park, J. Y. (2011). Effect of virtual reality on cognition in stroke patients. *Ann Rehabil Med*, *35*(4), 450-459. doi: 10.5535/arm.2011.35.4.450

Krch, D., Nikelshpur, O., Lavrador, S., Chiaravalloti, N. D., Koenig, S., & Rizzo, A. (2013). *Pilot results from a virtual reality executive function task*. Paper presented at the 2013 International Conference on Virtual Rehabilitation (ICVR).

Larson, E. B., Feigon, M., Gagliardo, P., & Dvorkin, A. Y. (2014). Virtual reality and cognitive rehabilitation: A review of current outcome research. *NeuroRehabilitation*, *34*(4), 759-772. doi: 10.3233/NRE-141078

Lustig, C., Tonev, C., & Hasher, L. (2000). *Visual distraction and processing speed I.* Paper presented at the 2000 Cognitive Aging Conference, Atlanta, GA.

Nombela, C., Nombela, M., Castell, P., Garcia, T., Lopez-Coronado, J., & Herrero, M. T. (2014). Alpha-theta effects associated with ageing during the Stroop test. *PLoS One*, *9*(5), e95657. doi: 10.1371/journal.pone.0095657

Parsons, T. (2011). Neuropsychological Assessment Using Virtual Environments: Enhanced Assessment Technology for Improved Ecological Validity. In S. Brahnam & L. C. Jain (Eds.), *Advanced Computational Intelligence Paradigms in Healthcare 6* (Vol. 337, pp. 271-289): Springer-Verlag Berlin Heiderberg.

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 Neuropsychol*, *13*(4), 363-381. doi: 10.1080/13825580600943473

Parsons, T. D., Courtney, C. G., Arizmendi, B., & Dawson, M. (2011). Virtual Reality Stroop Task for neurocognitive assessment. *Stud Health Technol Inform*, *163*, 433-439.

Rand, D., Weiss, P. L., & Katz, N. (2009). Training multitasking in a virtual supermarket: A novel intervention after stroke. *Am J Occup Ther*, *63*(5), 535-542.

Ratcliff, J. J., Greenspan, A. I., Goldstein, F. C., Stringer, A. Y., Bushnik, T., Hammond, F. M.,... Wright, D. W. (2007). Gender and traumatic brain injury: Do the sexes fare differently? *Brain Inj*, *21*(10), 1023-1030. doi: 10.1080/02699050701633072

Rizzo, A., & Kim, G. J. (2005). A SWOT Analysis of the field of VR rehabiliation and therapy. *Presence: Teleoperators and Virtual Environments*, *14*, 119-146.

Salisbury, D. B., Dahdah, M., Driver, S., Parsons, T. D., & Richter, K. M. (2016). Virtual reality and brain computer interface in neurorehabilitation. *Proc (Bayl Univ Med Cent)*, *29*(2), 124-127.

Schopp, L. H., Shigaki, C. L., Johnstone, B., & Kirkpatrick, H. A. (2001). Gender differences in cognitive and emotional adjustment to traumatic brain injury. *J Clin Psychol Med Settings*, *8*(3), 181-188.

Shallice, T., & Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain*, *114 (Pt 2)*, 727-741.

Simone, L. K., Schultheis, M. T., Rebimbas, J., & Millis, S. R. (2006). Head-mounted displays for clinical virtual reality applications: Pitfalls in understanding user behavior while using technology. *Cyberpsychol Behav*, *9*(5), 591-602. doi: 10.1089/cpb.2006.9.591

Appendix 1
Schedule of Intervention Administration

Session 1	Session 2 and 3	Session 4 and 5	Session 6 and 7	Session 8
VR Apartment Stroop (Baseline)	VR Apartment Stroop	VR Apartment Stroop	VR Apartment Stroop	VR Apartment Stroop
Condition 1: Inhibition task - Vigilance, successive discrimination, sustained attention task	Condition 1: Inhibition task - Vigilance, successive discrimination, sustained attention task	Condition 1: Inhibition task - Vigilance, successive discrimination, sustained attention task	Condition 1: Inhibition task - Vigilance, successive discrimination, sustained attention task	Condition 1: Inhibition task - Vigilance, successive discrimination, sustained attention task
Condition 2: Cognitive interference task – Sustained, selective, and divided attention task	Condition 2: Cognitive interference task – Sustained, selective, and divided attention task	Condition 2: Cognitive interference task – Sustained, selective, and divided attention task	Condition 2: Cognitive interference task – Sustained, selective, and divided attention task	Condition 2: Cognitive interference task – Sustained, selective, and divided attention task
External audio-visual distractions introduced for this Condition	No distractions introduced for both conditions	External auditory distractions introduced for both conditions	External visual distractions introduced for both conditions	External audio-visual distractions introduced for this Condition
VR Classroom Stroop (Baseline)	VR Classroom Stroop	VR Classroom Stroop	VR Classroom Stroop	VR Classroom Stroop
Condition 1: Inhibition task - Vigilance, successive discrimination, sustained attention task	Condition 1: Inhibition task - Vigilance, successive discrimination, sustained attention task	Condition 1: Inhibition task - Vigilance, successive discrimination, sustained attention task	Condition 1: Inhibition task - Vigilance, successive discrimination, sustained attention task	Condition 1: Inhibition task - Vigilance, successive discrimination, sustained attention task
Condition 2: Cognitive interference task – Sustained, selective, and divided attention task	Condition 2: Cognitive interference task – Sustained, selective, and divided attention task	Condition 2: Cognitive interference task – Sustained, selective, and divided attention task	Condition 2: Cognitive interference task – Sustained, selective, and divided attention task	Condition 2: Cognitive interference task – Sustained, selective, and divided attention task
External audio-visual distractions introduced for this Condition	No distractions introduced for both conditions	External auditory distractions introduced for both conditions	External visual distractions introduced for both conditions	External audio-visual distractions introduced for this Condition
ANAM; D-KEFS Stroop; WJ-III: Pair Cancellation; SSQ	———————	———————	———————	ANAM; D-KEFS Stroop; WJ-III: Pair Cancellation; SSQ
Speech Therapy Patients will receive training 5 days per week in attention, memory, and executive functioning techniques and tasks to measure functional improvements	Speech Therapy & Occupational Therapy Same as week 1, with complexity of tasks contingent on severity of patient's neurologic dysfunction	Speech Therapy & Occupational Therapy Same as week 1, with complexity of tasks contingent on severity of patient's neurologic dysfunction	Speech Therapy & Occupational Therapy Same as week 1, with complexity of tasks contingent on severity of patient's neurologic dysfunction	Speech Therapy & Occupational Therapy Same as week 1, with complexity of tasks contingent on severity of patient's neurologic dysfunction
Occupational Therapy Patients will receive training 5 days per week in paper-and-pencil tasks (e.g. clerical math; visuospatial tasks (pre-driving skills; parquetry), and/or IADLs to measure functional improvements				
Neuropsychological evaluation/ neurobehavioral status examination will occur at any point during patient's rehabilitation course - based on patient's ability to tolerate lengthy testing and medical status				