*****POSTPRINT: This is the final revised version of a manuscript that was published in *Neuropsychological Rehabilitation* (2014), 24 (2), 238-265, <http://dx.doi.org/10.1080/09602011.2014.881746>*****

Evaluation of a Virtual Reality Prospective Memory Task for use with Individuals with
Severe Traumatic Brain Injury

Short title: Prospective Memory in TBI

Allana L. Canty

Behavioural Basis of Health Research Program,
Griffith Health Institute and School of Applied Psychology, Griffith University

Jennifer Fleming

School of Health and Rehabilitation Sciences
Princess Alexandra Hospital & the University of Queensland

Freyr Patterson

Occupational Therapy Department
Princess Alexandra Hospital

Heather J. Green

Behavioural Basis of Health Research Program,
Griffith Health Institute and School of Applied Psychology, Griffith University

David Man

Department of Rehabilitation Sciences,
Hong Kong Polytechnic University

David H. K. Shum

Behavioural Basis of Health Research Program,
Griffith Health Institute, Griffith University

Correspondence: Professor David Shum, Griffith University, Gold Coast Campus, Queensland, 4222. Australia. E: d.shum@griffith.edu.au

Abstract

Objective: The current study aimed to evaluate the sensitivity, convergent validity and ecological validity of a newly developed virtual reality prospective memory (PM) task (i.e., the Virtual Reality Shopping Task; VRST) for use with individuals with traumatic brain injury (TBI). *Method:* Thirty individuals with severe TBI and 24 uninjured adults matched on age, gender and education level were administered the VRST, a lexical decision PM task (LDPMT), an index of task-friendliness and a cognitive assessment battery. Significant others rated disruptions in the TBI participants' occupational activities, interpersonal relationships and independent living skills. *Results:* The performance of the TBI group was significantly poorer than that of controls on event-based PM as measured by the LDPMT, and on time- and event-based PM as measured by the VRST. Performance on the VRST significantly predicted significant others' ratings of patients' occupational activities and independent living skills. The VRST was rated as significantly more reflective of an everyday activity, interesting and was afforded a higher recommendation than the LDPMT. For the TBI group, event- and total PM performance on the VRST significantly correlated with performance on measures of mental flexibility and verbal fluency, and total PM performance correlated with verbal memory. *Conclusion:* These results provide preliminary but promising evidence of the sensitivity, as well as the convergent and ecological validity of the VRST.

Key words: Prospective memory, virtual reality, brain injury, ecological validity, assessment.

Evaluation of a Virtual Reality Prospective Memory Task for use with Individuals with Severe Traumatic Brain Injury

Memory failures persist in about one quarter of individuals who sustain traumatic brain injury (TBI) and are often the source of chronic functional disability (Goldstein & Levin, 1995; Levin & Hanten, 2002). One significant challenge for persons with TBI reintegrating themselves into the community is the inability to recall delayed intentions. This type of memory failure is referred to as prospective memory (PM) impairment. PM is instrumental to optimal functioning as a number of everyday tasks depend on the integrity of this function (e.g., following a medicine regime and the timely keeping of appointments) (Schmitter-Edgecombe, Greeley, & Woo, 2009; Woods et al., 2009).

PM has typically been classified according to the nature of the cue that triggers retrieval of the delayed intention (McDaniel & Einstein, 2007). Time-based PM involves carrying out an intended action at a certain time or after a period of time has lapsed (e.g., picking up children at 3pm). A PM task is considered event-based when the cue to execute an intention is an environmental signal (e.g., oven alarm signals to remove muffins). Lastly, activity-based PM tasks are characterised by the need to carry out an intended action at the end of an activity or before starting a new activity (e.g., buying milk after work).

Shum, Levin, and Chan (2011) used meta-analysis to summarise those studies which have provided evidence of PM deficits in the TBI population. Among studies that used self- or other-report questionnaires, results indicated that individuals with TBI and their significant others frequently report PM deficits. Furthermore, there was consistent evidence from behavioural studies suggesting individuals with TBI are impaired on both time- and event-based PM (Maujean, Shum, & McQueen, 2003; Shum, Valentine, & Cutmore, 1999) and that prospective forgetting is linked with deficits in attention (Mathias & Mansfield, 2005), executive functioning (e.g., mental flexibility and working memory; Fleming et al., 2008;

Maujean et al., 2003; Knight, Harnett, & Titov, 2005), and verbal memory (Schmitter-Edgecombe & Wright, 2004). Despite consistent evidence of PM deficits in the TBI population, evidence of the reliability, sensitivity and validity (particularly ecological validity) of the measures used to acquire such results is limited.

Evaluating the ecological validity of neuropsychological tests has been increasingly recognised as important over the past decade. Chaytor and Schmitter-Edgecombe (2003) conceptualised a two-tier definition of ecological validity: verisimilitude and veridicality. Verisimilitude is the degree to which the cognitive demands of a test theoretically and topographically resemble the cognitive investment required by a task performed in one's everyday environment. The related concept of veridicality refers to the degree to which a test predicts everyday functioning. This heightened emphasis on ecological validity has paralleled a time in which many neuropsychologists have unearthed the limitations of many conventional tests of memory and are seeking new approaches to measure cognitive impairments. As such, the benefits associated with virtual assessment platforms have attracted considerable attention (Knight & Titov, 2009).

Virtual reality (VR) is a term used to describe computer-generated artificial environments with distinctive sensory properties that can be explored and interacted with in real time (Knight & Titov, 2009). This innovative medium of psychological assessment has the potential to balance the demands of ecological validity with the sensitivity and specificity of conventional measures of PM and overcome the pragmatic and reliability issues associated with functional measures (e.g., The Multiple Errands Task, Shallice & Burgess, 1991; The Real Library Task, Renison, Ponsford, Testa, Richardson, & Brownfield, 2012). A range of VR environments have been used to assess time-, event- and activity-based PM (e.g., Jansari, Agnew, Akesson & Murphy, 2004; Sweeney, Kersel, Morris, Manly, & Evans, 2010) and executive functioning (including PM tasks), including a virtual supermarket (Kinsella, Ong, & Tucker, 2009; Klinger,

Chemin, Lebreton, & Marie, 2006), virtual breakfast task (Craik & Bialystok, 2006), virtual university (McGeorge et al., 2001), virtual office (Jansari et al., 2004; Logie, Trawley, & Law, 2011), virtual beach (Elkind, Rubin, Rosenthal, Skoff, & Prather, 2001), and virtual street (Titov & Knight, 2009). Recently a VR version of the Multiple Errands Task has been developed, but the utility of this program has only been demonstrated in a group of nine post-stroke patients (Rand et al., 2009). Research on these tools has been largely exploratory and has lacked psychometric rigor (Knight & Titov, 2009). Validation of the relationship between VR and real life environments and their utility within TBI populations has seldom been addressed.

One notable exception is the study performed by Renison et al. (2012) which investigated the ecological validity and construct validity of a newly developed VR measure of EF, the Virtual Library Task (VLT). This measure included tasks assessing time and event-based PM. In addition to the VLT, participants ($n = 30$ TBI and $n = 30$ healthy controls) were administered a real life analogous task - the Real Library Task (RLT) and five neuropsychological measures of EF. Significant others for each participant also completed the Dysexecutive Questionnaire (DEX), which is a behavioural rating scale of everyday EF. Performances on the VLT and the RLT were significantly positively correlated indicating that VR performance is similar to real world performance (e.g., time based PM = 5.48, $p < .01$; event-based PM $r = 5.73$, $p < .01$). The TBI group performed significantly worse than the control group on the VLT and the Modified Six Elements Test (MSET) but the other four neuropsychological measures of EF failed to differentiate the groups. Both the MSET and the VLT significantly predicted everyday EF suggesting that they are both ecologically valid tools for the assessment of EF. The VLT has the advantage over the MSET of providing objective measurement of individual components of executive functioning.

Potvin, Rouleau, Audy, Charbonneau, and Giguere (2011) sought to evaluate the validity and sensitivity of another virtual task - the Test Ecologique de Mémoire Prospective (TEMP) - with individuals with moderate and severe TBI ($n = 30$). The TEMP involves a 20-minute movie that presents the graphical perspective of a driver navigating a city. Test-takers are required to perform 10 event-based and five time-based tasks to prepare for a birthday dinner (version A) or a holiday (version B). Results demonstrated that individuals with TBI experienced more pronounced problems than matched controls ($n = 15$) in learning the content of delayed intentions (retrospective component) and retrieving these intentions in the correct context (prospective component), especially in the time-based condition. Correlational results suggested that the retrospective component of PM was mainly supported by episodic memory processes, while the prospective component was predominantly supported by processes involved in attention and executive functioning. Moreover, there was a significant correlation between performance on the TEMP and results on a questionnaire assessing PM functioning in daily living completed by participants' relatives. Taken together, these results provide preliminary evidence that the TEMP is a sensitive tool for assessing PM deficits in the TBI population (for other virtual reality studies of PM, see Knight et al., 2005; Sweeney et al., 2010).

Although the aforementioned studies successfully demonstrated the utility of virtual assessment platforms, a number of adjustments are recommended to enhance the endorsement of virtual assessment within clinical settings and the experimental literature. First, with the exception of Sweeney et al. (2010) and Renison et al. (2012), the level of sensitivity to TBI deficits and validity that VR programs offer relative to conventional experimental measures of PM has seldom been explored. Furthermore, the measures discussed do not record time monitoring frequency, thus cannot be used to derive inferences about the reasons underpinning time-based PM failures. Finally, many studies (e.g., Knight et al., 2005; Potvin et al., 2011)

have not compared the relative veridicality and verisimilitude of their virtual platforms to that of conventional experimental measures of PM. Accordingly, conclusions concerning the ecological validity of such measures should be drawn tentatively. These limitations were addressed in the present study.

The first objective of the present study was to develop a VR test of PM (i.e., the Virtual Reality Shopping Task; VRST) that is sensitive to PM dysfunction and addresses the limitations of previous measures. Principally, the VRST was designed to be more exhaustive than previous VR and conventional computerised PM instruments as it measures time- and event-based PM concomitantly and also measures time-monitoring frequency. The second objective was to demonstrate the added ecological validity (i.e., veridicality and verisimilitude) of the VRST over that of a conventional computerised assessment of event-based PM (i.e., a lexical decision prospective memory task; LDPMT). Third, the study sought to evaluate the convergent and discriminant validity of the VRST by exploring the relationships between PM performance and measures of cognitive functioning. Specifically, it was hypothesised that: (1) Participants with TBI would perform significantly poorer than non-injured adults on time- and event-based PM as measured by the VRST and on event-based PM as measured by the LDPMT; (2) The sensitivity of the VRST would be comparable to, if not better than, that of the LDPMT; (3) For the TBI group, PM performance as measured by the VRST would correlate significantly with indices of psychosocial functioning (i.e., disruptions of occupational activities, interpersonal relationships and independent living skills); (4) Participants would rate the VRST as significantly more reflective of an everyday activity, as more interesting and would afford the task a higher recommendation than the LDPMT, and (5) PM performance as measured by the VRST would correlate strongly with measures of cognitive functioning (i.e., attention, mental flexibility, verbal fluency, verbal memory, response initiation/inhibition and working memory).

Method

Participants

Thirty participants with severe TBI (25 males, 5 females) were recruited from the Brain Injury Rehabilitation Unit at Princess Alexandra Hospital in Brisbane, Australia. Patients were recruited on a voluntary basis according to the following inclusion criteria: (1) Aged between 18 and 60 years, (2) initial Glasgow Coma Scale (GCS) score between three and eight, (3) emerged from a period of post traumatic amnesia (PTA; greater than 24 hours), (4) less than two years post injury, (5) literate in English, (6) ambulant or independently mobile in a manual or electric wheelchair, (7) cognitive capacity to provide informed consent and, (8) have a significant other available to participate in the study. Cause of injury varied and consisted of 12 motor vehicle accidents, seven assaults, six falls, and five motorbike accidents. Exclusion criteria included prior PM training, a known psychiatric diagnosis, a prior incidence of TBI or a history of other neurological disorders (viz., stroke, hypoxic injury or neurodegenerative disorder), substance abuse, learning disabilities (including premorbid intellectual disability) and physical or cognitive inability to use a computer. TBI participants were screened for these conditions via consultation with their treating clinician and examination of their medical record.

A comparison group of 24 healthy adults (20 males and 4 females) with no history of a significant neurological event, psychiatric illness and drug or alcohol dependency (assessed via a brief structured interview) were recruited from the general community using the first author's personal contacts. The sample sizes of the TBI and control groups were determined by the availability of TBI participants who consented to take part in the study during the recruitment period and are comparable to other studies of PM and virtual reality in TBI (e.g., Potvin et al., 2011; Renison et al., 2012).

All participants were Caucasian and TBI and healthy participants were equivalent in age ($t(54) = -.66, p = .51$), years of education ($t(54) = 1.68, p = .10$), and in the proportion of males and females ($\chi^2(1) = .56, p = .72$). These and other socio-demographic and clinical characteristics are summarised in Table 1. The study received ethical approval from the Human Research Ethics Committees of Griffith University and the Princess Alexandra Hospital. Informed consent was obtained from all participants in accordance with the National Statement on Ethical Conduct in Research Involving Humans (NHMRC, 2007).

Insert Table 1

Measures

The VRST is a measure of time- and event-based PM modelled off the classical dual-task paradigm of PM (Einstein & McDaniel, 1990) and Ellis' (1996) theoretical stages of PM. The ongoing task simulates the graphical perspective of the participant's avatar and requires the participant to purchase 12 items in a pre-specified order from a selection of 20 shops (see Figure 1). This list can be retrieved for checking at any time. In addition, participants were instructed to complete three time- and three event-based PM tasks. The time-based tasks required test-takers to access a virtual mobile phone to send text messages at the 4th, 8th and 12th minute of the task. Participants were not required to recall or type message content. Rather, participants were instructed that pre-written content (e.g., "Happy Birthday") was automatically sent by clicking the green send button on the virtual phone (see Figure 2). Participants were permitted to check the time on the screen as often as necessary. The event-based task required test-takers to press the 'T' key on the computer's keyboard when an audio 'sale' announcement was heard. Three sale announcements, occurring at the 3rd, 9th and 13th minute, are presented amidst three non-sale announcements, which occur at the 5th, 7th and 11th

minute of the task. The total duration of the task is 14 minutes. Responses to PM trials were scored according to the deviation from the target time (Einstein, McDaniel, Richardson, Guynn, & Cunfer, 1995; Shum et al., 1999; e.g., 3 points = +/- 10s; 2 points = + 11-20s; 1 point = 21-40s)¹. Time-based scores allowed for responding before or after the target time, whereas event-based scores only allowed responding after the sale announcement had been presented. The program was written using Virtools and run through 3DVIA-player on a DELL laptop computer (E6510).

Insert Figure 1

Insert Figure 2

The *LDPM*T, also an adaption of Einstein and McDaniel's (1990) dual-task paradigm, provides a brief measure of event-based PM. The ongoing task requires participants to indicate whether presented stimuli are words or non-words. Words and non-words were generated using the MRC Psycholinguistic Database (Coltheart, 1981) and ARC Nonword Database (Rastle, Harrington, & Coltheart, 2002) respectively. The parameters used to generate stimuli were selected on the basis of those used in previous research² (Einstein & McDaniel, 2005; Marsh, Hicks, & Bink, 1998; Smith, 2003). The PM component required participants to press the 'SPACE' bar on the computer's keyboard when a PM cue (i.e., an animal word) was detected. PM cues were presented at the 75th, 210th, 315th, 385th, 460th and 530th second of the task. The task consisted of 126 trials (60 words; 60 non-words; 6 animal words) which required 10 minutes and 30 seconds to complete. Participants were given a score of 1 to 6 depending on the number of PM cues they correctly detected (i.e., higher scores are indicative of better

performance). A PM response was recorded as correct if the participant responded to the cue (i.e., before or after responding to the ongoing task) within the allocated 3 sec performance interval (Ellis, 1996). Tasks of this nature are theoretically derived and have previously been used to assess PM in the TBI literature, thus are a good criterion against which to validate the VRST (e.g., Maujean, Shum, & McQueen, 2003; McCauley et al., 2011; Shum et al., 1999).

Cognitive Measures. Participants were administered a battery of cognitive tests commonly used in clinical and experimental settings to evaluate attention and mental flexibility (Trail Making Test: Parts A and B; Reitan & Wolfson, 1985), verbal memory (Hopkins Verbal Learning Test-Revised; HVLTR; Brandt, 1991), verbal fluency (Controlled Oral Word Association Task; COWAT; Benton & Hamsher, 1978), response initiation/inhibition (Hayling Test; Burgess & Shallice, 1997) and working memory (Letter Number Sequencing Subtest of the Wechsler Memory Scale-III; LNS; Wechsler, 1997). As PM is closely related to the processes assessed by these measures (Fleming et al., 2008; Knight et al., 2005; Renison et al., 2012), they were selected to provide evidence of convergent validity.

Self-Report Questionnaires. To assess group differences in the quality of participants' testing experience and to quantify the psychosocial functioning of TBI patients, the following measures were administered: (a) *The User-Friendliness Scale* (UFS; see Appendix A) is a 15-item self-report questionnaire developed so that the researchers could obtain participants' objective ratings of the degree to which the VRST and the LDPMT reflected an everyday activity, were interesting, clear and easy to learn, and the extent to which they would recommend each task. The items included in this scale were modelled off Chaytor and Schmitter-Edgecombe (2003) definition of ecological validity; (b) *The Sydney Psychosocial Reintegration Scale (SPRS) Form B* (Tate, Hodgkinson, Veerabangsa, & Maggiorotto, 1999) is a 12-item other-report questionnaire that quantifies disability in three domains of everyday living commonly disrupted after TBI: occupational activities, interpersonal relationships and

independent living skills.

Procedures

Participants were administered the tasks in the following order: the two PM tasks, the battery of cognitive tests (using standard procedures), and the UFS. The order in which the PM tasks were administered was counterbalanced to control for testing effects. To ensure participants understood the requirements of the PM tasks, participants were asked to free-recall their respective instructions. Similarly, at the end of each task, participants were again asked to describe the instructions associated with the measure. Such information permitted the experimenter to differentiate a participant's impairment on the retrospective component of the task from that on the prospective component. Close others of TBI participants were asked to complete the SPRS. The total administration time was approximately two hours, with some participants tested over two sessions if they were vulnerable to fatigue.

Data Analysis

Demographic comparisons between patients and controls were conducted using a Chi Square test for gender and independent samples *t*-tests for age and years of education. Group differences in PM performance were assessed using a one-way analysis of variance (ANOVA) for LDPM performance, and factorial ANOVA (Group x Task) for VRST performance. Independent-samples *t*-tests were used for post-hoc testing of group differences in time and event-based PM performance. Pearson's correlations were performed to assess the relationships between PM performance, executive functioning and psychosocial functioning. Sensitivity and specificity as well as positive and negative predictive values for VRST time-, event- and total-PM were calculated. Alike Sweeney et al. (2010), the absence of large-scale norms precluded the definition of impairment. As such, selection of "cut-off" scores to examine sensitivity/specificity was necessarily exploratory. Standard multiple regressions were performed to examine the relative contribution of PM measures and cognitive abilities to

predict psychosocial functioning as measured by the SPRS. Although the current sample size is less than optimal for regression analyses (Tabachnick & Fidell, 2007), preliminary exploration of these relationships was considered interesting and useful given that not much is known about them. Caution will be taken in interpreting the results of such analyses in the Discussion.

Results

LDPMT. Table 2 summarises the performance of the TBI group and controls on the LDPMT. The TBI group correctly responded to significantly fewer words ($t(42.27) = 5.28, p < .001, d = 1.07$), non-words ($t(35.73) = 4.07, p < .001, d = 1.04$) and animal words ($t(52.11) = 4.41, p < .001, d = 1.37$) than did controls.

VRST. The TBI group performed significantly poorer than controls on the ongoing component of the VRST, $t(49.44) = 3.54, p = .001, d = .92$ (see Table 2). A 2 (Group) x 2 (PM Task) ANOVA yielded significant main effects for PM Task ($F(1, 54) = 49.75, p < .001, \eta_p^2 = .48$) and Group ($F(1, 54) = 44.67, p < .001, \eta_p^2 = .45$). Furthermore, the analysis yielded a significant Group x PM Task interaction, $F(1, 54) = 7.29, p = .009, \eta_p^2 = .12$. Both groups performed significantly poorer on the time- than the event-based PM tasks (controls = $t(24) = -3.62, p = .001, d = 1.01$; TBI = $t(30) = -6.45, p < .001, d = 1.49$). Between group analyses indicated that the TBI group performed significantly poorer than controls on both the time- ($t(54) = 8.03, p < .001, d = 2.17$) and the event-based PM component of the VRST ($t(45.97) = 3.09, p = .003, d = .83$). Interestingly, the TBI group did not make significantly more commission errors ($M = .32, SD = .83$) than controls ($M = .28, SD = .61$), $t(54) = -.21, p = .83$. Furthermore, all participants demonstrated accurate post-assessment recall of the instructions accompanying the LDPMT and VRST. This suggests that prospective forgetting was indicative of poor performance on the prospective, rather than the retrospective, component of the task. Taken together, such results indicate that the interaction arose because the TBI group was significantly more impaired on the time- than the event-based PM tasks compared to controls.

Insert Table 2

Time Checking Behaviour. A 2 (Group) x 4 (Interval) mixed ANOVA was performed to assess the effect of Group and Interval (i.e., the 1st, 2nd, 3rd and 4th minute of each 4 minute block preceding a time based response³) on time monitoring frequency (see Table 2). Significant main effects were observed for Interval ($F(3, 162) = 27.09, p < .001, \eta_p^2 = .33$) and Group ($F(1, 54) = 44.02, p < .001, \eta_p^2 = .44$). Additionally, a significant Group x Interval interaction was identified, $F(3, 162) = 14.29, p < .001, \eta_p^2 = .22$. Controls checked the time significantly more frequently than the TBI group within the 3rd ($t(54) = 5.33, p < .001, d = 1.42$) and 4th interval ($t(54) = 6.47, p < .001, d = 1.71$). Two separate repeated measures ANOVAs were performed to assess within-group differences in time monitoring frequency. Although time monitoring frequency increased significantly across the four intervals for controls ($F(2.67, 54.38) = 29.57, p < .001, \eta_p^2 = .55$), this was not the case for the TBI group ($F(3, 90) = 1.79, p = .15$). Paired samples *t*-tests were conducted to identify the intervals between which significant differences in time monitoring were observed for controls ($\alpha = .004$). Significant differences in time monitoring were identified between the 1st and the 3rd interval ($t(24) = -4.06, p < .001, d = 1.24$), 1st and 4th interval ($t(24) = -6.22, p < .001, d = 1.98$), 2nd and 3rd interval ($t(24) = -4.77, p < .001, d = 1.14$), 2nd and 4th interval ($t(24) = -8.50, p < .001, d = 1.91$) and the 3rd and 4th interval ($t(24) = -4.60, p < .001, d = .95$). Taken together, the results indicate that the interaction arose because controls, in contrast to the TBI group, checked the time significantly more frequently in the intervals immediately preceding each time based task. Medium to large correlations (two-tailed) were observed between VR time-based PM performance and time monitoring frequency in the 3rd interval ($r = .41, p < .05$) for controls and in the 3rd ($r = .48, p = .006$) and 4th interval ($r = .57, p < .001$) for the TBI group.

Relationship between VRST and LDPMT. Correlational analyses were conducted to explore the relationship between performance on the prospective components of the VRST and the LDPMT

(see Table 3). Significant and moderate correlations were found between performance scores on the event-based PM components of the LDPMT and VRST for the TBI group and controls and between total PM performance on the VRST and event-based PM performance on the LDPMT for the TBI group. No significant correlations were observed between event-based LDPMT performance and time-based PM performance on the VRST.

Sensitivity and Specificity of VRST and LDPMT. The reference criterion for the sensitivity and specificity calculations was group membership (i.e., TBI vs healthy control). The spread of scores for the two groups were examined and those scores which maximised both sensitivity and specificity were selected. For the VRST time-based PM score, a cut-off score of five or below was selected to represent impairment. With this score, results indicated a sensitivity of 68%, specificity of 97%, positive predictive value of 94% and a negative predictive value of 70%. For the VRST event-based PM score, a score of six or below was selected as representing impairment. Results indicated a sensitivity of 76%, specificity of 71%, and positive predictive value of 79% and negative predictive value of 65%. For the VRST total PM score, a score of 12 or below was selected as representing impairment. Results indicated a sensitivity of 76%, specificity of 90%, positive predictive value of 88% and negative predictive value of 74%. For the LDPMT event-based PM score, a score of four or below was selected as representing impairment. Results indicated a sensitivity of 96%, specificity of 69%, positive predictive value of 83% and negative predictive value of 57%.

With the above cut-off values it can be seen that while sensitivity and specificity is high for VRST time and total PM, the specificity of the event-based PM components of the VRST and LDPMT are modest in that only just over half of the TBI group scored at or below the cut-off on both core measures. Given the fact that prospective remembering reflects a number of different cognitive skills it was possible that some of the patients referred to the study either

did not actually have PM deficits or that they had deficits in aspects of PM not captured by the VRST (e.g., activity-based PM).

Cognitive Measures. Table 2 summarises the groups' performance on the cognitive measures. The TBI group performed significantly poorer than controls on measures of attention and mental flexibility (Trails A, $t(35.66) = 5.70, p < .001, d = 1.46$; Trails B, $t(35.25) = 7.11, p < .001, d = 1.82$), verbal memory (HVL-R total recall, $t(54) = 5.33, p < .001, d = 1.45$), verbal fluency (COWAT, $t(54) = 4.10, p < .001, d = 1.12$), response initiation/inhibition (HSCT, $t(40.18) = 4.94, p < .001, d = .67$) and working memory (LNS, $t(54) = 5.82, p < .001, d = 1.55$).

Relationship between Cognitive Tests and PM Performance. Table 3 summarises the correlations between PM performance and cognitive functioning for the TBI group and controls. For the TBI group, significant relationships were found between VRST event-based PM and measures of mental flexibility and verbal fluency and between total VRST PM performance and mental flexibility (i.e., Trails B), verbal memory and verbal fluency. Further correlations were observed between event-based PM as measured by the LDPMT and verbal memory and verbal fluency. Although not significant, several small to moderate correlations were found between the TBI group's PM performance as measured by the VRST and LDPMT and measures of attention, response initiation/suppression and working memory. Similarly, for controls, several small and medium correlations (albeit not significant) were found between measures of cognitive functioning and PM performance, as measured by the VRST and LDPMT.

Relationship between Psychosocial Functioning, PM Performance and Cognitive Measures. As seen in Table 3, significant and moderate correlations were observed between time-, event- and total PM performance as measured by the VRST and Independent Living Skills as measured by the SPRS. Additionally, a significant moderate correlation was observed between time-based PM performance on the VRST and other's ratings of Occupational Activities. Although not significant, small to medium correlations were observed between event-based PM performance as measured by the LDPMT and measures of psychosocial functioning. Furthermore, significant

correlations were observed between measures of mental flexibility and verbal fluency and the occupational activities and independent living subscales of the SPRS for the TBI group (see Table 3).

Two standard multiple regression analysis were conducted to examine the relative contributions of PM and cognitive ability to scores on the occupational activities and independent living subscales of the SPRS. Four predictor variables were included in each regression, on the basis that they were correlated with the SPRS: VRST time-based PM, VRST event-based PM, Trails B and the COWAT. Variables were entered simultaneously into the regression equations to ascertain the unique contribution of each to variance in SPRS scores. A summary of the results of the regression analyses are presented in Table 4. For the occupational activities subscale, the total model was significant, $F(1, 28) = 15.94, p < .001$, explaining 36.3% of the variance in SPRS scores. At an individual level, time-based PM and Trails B were significant, with the Beta values indicating that time-based PM was the stronger predictor of scores on the occupational activities subscale of the SPRS.

For the independent living subscale, the total model was also significant, $F(1, 28) = 6.31, p < .05$, explaining 18.4% of the variance in SPRS scores. At an individual level, time- and event-based PM were significant, with the Beta values indicating that time-based PM was the stronger predictor of scores on the independent living subscale of the SPRS.

User-Friendliness Scale. Table 5 summarises the mean ratings provided by the TBI group and controls on each item of the UFS. A series of 2 (Group) x 2 (PM Task) mixed ANOVAs were conducted to identify differences in how the TBI group and controls rated the user-friendliness of the VRST and LDPMT (see Table 5). Where interactions were observed, simple main effects were explored by way of independent and paired samples *t*-tests. Examination of the means suggests: participants rated the VRST as more reflective of an everyday activity and as more interesting than the LDPMT; the TBI group, relative to controls, found the VRST and LDPMT to be slightly more reflective of everyday activities and rated the tasks as slightly more difficult to understand.

Interestingly, the extent to which participants recommended the VRST far exceeded the extent to which they recommended the LDPMT.

Insert Table 3

Insert Table 4

Insert Table 5

Discussion

Preliminary evidence suggests the VRST has sound convergent and ecological validity and is a sensitive measure of PM dysfunction following TBI, thereby achieving the study's primary objective. Hypothesis one, that participants with TBI would perform significantly poorer than controls on time- and event-based PM as measured by the VRST and on event-based PM as measured by the LDPMT, was supported by the results. An additional finding was that participants performed significantly poorer on the time- than the event-based PM component of the VRST. This finding converges with Einstein et al.'s (1995) theorisation that time-based PM tasks are more difficult than event-based tasks because the former are more reliant on self-initiated retrieval and subjective time estimation than the latter. Prior to the present study, Einstein et al.'s (1995) supposition was not empirically supported (see Shum et al., 2011). By embedding two subtypes of PM in the one ongoing task, rating performance along a comparable scale and using a within-subject design, the current study was able to compare the relative difficulty of time- and event-based PM tasks and provide some evidence that the former is more difficult than the latter.

Consistent with the test-wait-test-exit model (TWTE; Harris & Wilkins, 1982), group differences in VR time-based PM performance corresponded with the frequency with which the

groups' strategically monitored the time. Compared to the TBI group who did not monitor the time significantly more frequently as the task progressed, controls monitored the time significantly more often in periods immediately before the target time (i.e., 3rd and 4th interval) than in earlier periods (i.e., 1st and 2nd interval). Notwithstanding, the performance of the TBI group on the VR time-based task correlated significantly with monitoring frequency in intervals (3rd and 4th) proximal to the target time. Such findings provide corroborating evidence for the TWTE model which indicates that time-monitoring is an important variable for successful performance on time-based PM tasks (Shum et al., 1999). For controls however, VR time-based PM performance was only significantly related to time monitoring in the 3rd interval. This finding is atypical and suggests that greater monitoring during the middle period of the task was an important determinant for the shorter response latencies of controls. One plausible explanation of this result is that controls in the present study used the 3rd interval as an anchor when responding to the target time. This anchor may have allowed participants to judge the passing of time and as a result, respond closer to the target time. This interpretation is speculative, and highlights the need for further research into the effect of strategy use on time-based PM performance.

The TBI group did not make significantly more commission errors than controls. While this may be contrary to expectation, this result is consistent with the finding that TBI participants' VR event-based PM performance did not significantly correlate with their performance on a measure of response initiation and suppression (i.e., the HSCT-R). Examination of the raw data suggests that the TBI group either neglected to respond to PM cues or responded too slowly. Taken together, this pattern of results suggests that successful event-based PM is a function of self-initiation processes rather than response inhibition. Furthermore and consistent with Potvin et al. (2011), because the TBI group performed significantly poorer than controls on the ongoing component of both the VRST and LDPMPT it can be concluded that group differences in PM performance did not arise because the TBI group, relative to controls, failed to share their cognitive resources between the ongoing and PM task components. Similarly, all participants included in the analyses accurately

recalled the intentions associated with the two PM tasks post-assessment. Thus prospective forgetting was likely due to impaired performance on the prospective rather than the retrospective component of the task in question.

Preliminary support was found for hypothesis two that the sensitivity of the VRST would be comparable to, if not better than, that of the LDPMT. As expected, the effect sizes accompanying group differences in PM performance as measured by the VRST and LDPMT were large (Cohen, 1988) and the relative classification accuracy of each measure was modest. These results suggest that the LDPMT and VRST are sensitive to PM deficits in the TBI population and provide preliminary evidence that these measures differentiate individuals with TBI from healthy controls. This study is the first to suggest that performance on virtual measures of PM are similar to performance on conventional experimental measures of PM. This finding is important because virtual assessment has greater clinical utility than assessment in real-world settings and has greater veridicality than conventional measures of PM. This finding therefore supports the use of virtual environments for testing of patients with TBI in experimental settings. Confidence in the sensitivity of the VRST, and in turn its utility in clinical settings, needs to be improved by future research which demonstrates robust relationships between the VRST and more established clinical or functional measures of PM (e.g., Real Library Task; CAMPROMPT).

Evidence was also found to support the third hypothesis that for the TBI group PM performance as measured by the VRST would correlate significantly with indices of psychosocial functioning (i.e., disruptions of occupational activities, interpersonal relationships and independent living skills). Interestingly, unlike the VRST, the LDPMT was not significantly correlated with the domains of psychosocial functioning. However, the observed correlations were comparable between tasks, suggesting that the two measures offer a relatively similar level of veridicality. This aside, the finding that the TBI group's time-based, event-based and total PM performance on the VRST significantly correlated with significant others' perceptions of their everyday independent living skills supports the veridicality of the task and has important

implications for clinical practice. The finding suggests that it is the dysfunctional interplay of PM processes, in addition to either time- or event-based PM alone that threatens functional independence. Accordingly, the management of PM impairment may need to shift from treating the different subtypes of PM in isolation to focusing on the interplay of time- and event-based prospective remembering.

Furthermore, time-based PM was the strongest predictor of occupational functioning and time- and event-based PM were the strongest predictors of independent living skills, providing support for a significant association between performance on the VRST and psychosocial functioning as measured by the SPRS. Nevertheless there remains a substantial proportion of the unshared variance in SPRS scores, suggesting additional factors are contributing to the reported vocational and independent living difficulties. These may include: the level of environmental cognitive demand and compensatory strategy use (Renison et al., 2012), how often the person's PM impairments negatively affect the significant other (Joyner, Silver & Stavinocha, 2009), the degree of acceptance of cognitive and behavioural changes in the injured participant, emotional problems, pre-morbid functioning and health problems. Future research examining ecological validity should attempt to control for the above mentioned sources of variation where possible. Given the small sample size, results demonstrating relationships between PM impairment and psychosocial functioning in the TBI population should be regarded as preliminary, and require replication.

Evidence supporting hypothesis four that participants would rate the VRST as significantly more reflective of an everyday activity, as more interesting and would afford the task a higher recommendation than the LDPMT, suggests that the VRST offers greater verisimilitude (e.g., naturalistic stimuli and PM tasks) than the LDPMT. Taken together, findings concerning veridicality and verisimilitude (i.e., ecological validity) suggest that the VRST may be a useful supplementary clinical test to confirm and specify PM problems experienced by patients with neurological and psychiatric conditions known to disrupt PM. The clinical importance of face validity should be recognised; patients are more likely to accept feedback regarding cognitive

impairments if the tools used to make decisions regarding their level of deficit are reflective of scenarios in which they engage in their everyday environment.

Lastly, the study found preliminary evidence of convergent validity of the VRST (i.e., the study's third objective) by exploring the relationships between PM performance and measures of cognitive functioning. Exploring these relationships also provided partial support for the final hypothesis that PM performance as measured by the VRST would correlate strongly with measures of cognitive functioning. As expected, the TBI group's event- and total PM performance on the VRST significantly correlated with an index of mental flexibility. These findings are consonant with those of earlier studies that have demonstrated individuals who are unable to switch and focus attention are unlikely to monitor their environment in a way that facilitates PM cue detection (Mathias & Mansfield, 2005). Furthermore, consistent with the results of Knight et al. (2005), significant correlations were observed between PM (i.e., event- and total VRST performance and event-based LDPMT performance) and a measure of verbal fluency, which may reflect the verbal nature of the tasks and the fact that both tasks require self-generation of response strategies. Furthermore, the correlations between RM and PM task performance (i.e., total VRST performance and event-based LDPMT performance) is consistent with previous research (e.g., Schmitter-Edgecombe & Wright, 2004) and PM theory that PM incorporates a RM component (Burgess & Shallice, 1997; Einstein & McDaniel, 1990). The absence of significant correlations between VRST time-based PM performance and cognitive processes, and between VRST event-based PM performance and attention and working memory is unexpected in light of previous research (e.g., Shum et al., 1999) and should be addressed in further studies utilising the VRST.

Ancillary evidence of construct validity was suggested by the strong correlations observed between event-based and total PM scores on the VRST and performance on the LDPMT (i.e., convergent validity). Lexical decision tasks have strong theoretical architecture (viz., Einstein & McDaniel, 1990; Ellis, 1996) and have become the benchmark of PM assessment in the

experimental literature. As such, finding a relationship between the VRST and LDPMT provides preliminary support for the convergent validity of the former measure. It is important for future research to further explore the construct and criterion validity of the VRST by inspecting the strength of correlations between this task, self-report (e.g., Brief Assessment of Prospective Memory; Man, Fleming, Hohnsbein, & Shum, 2011), psychometric measures of PM (e.g., CAMPRM; Wilson et al., 2005) and functional measures of PM (Renison et al., 2012).

A number of limitations relevant to the present study must be acknowledged. One limitation was that the VR event-based PM performance of controls approached a ceiling effect. One possible explanation for this was that there was only a 3-minute delay between the encoding phase and the first event-based cue in the VRST. The challenge for future program developers will be to balance this delay period against designing short ecological event-based PM tasks which are sensitive to the spectrum of functioning. A possible solution is to have participants generate responses to conspicuous event-based cues (e.g., passing on a message to a friend's avatar). This would have the advantage of increasing the difficulty of the overall program, which could in turn make the task more sensitive to milder cognitive impairments.

Second, the sensitivity of the VRST could not be adequately explored, as a comparable experimental platform was not available. That is, the LDPMT did not contain a time-based PM component, nor did it include auditory event-based PM cues. To demonstrate the added sensitivity of the VRST over that offered by cheaper conventional experimental platforms, a more comparable computerised task needs to be developed. Such a program would include a time-based PM component and a broader range of event-based PM cues that differ in salience and tap different sensory modalities. Furthermore, PM ability as measured by the conventional experimental platform would need to be rated on a scale which is comparable to that used in the VRST (e.g., time deviation of response). Third *p* values were not corrected to account for the large number of computed correlations. The authors acknowledge that these findings are preliminary and require further replication before strong conclusions can be drawn. Fourth, as the sample size of the TBI

group is modest, results concerning the predictive value of the VRST, and the relative sensitivity and specificity of the VRST and LDPMT, although encouraging, are preliminary and require replication with a larger sample of individuals with TBI. Fifth, because the TBI group performed poorly on the ongoing component of the VRST and LDPMT, it is difficult to discern whether impaired performance on these tasks reflect a PM problem specifically, instead of a general effect of task difficulty or deficits in other cognitive domains. However, as the TBI group had all acquired severe injuries, it is not likely they would perform equivalent to the control group on the ongoing task. Furthermore, this pattern of result is comparable to that of other virtual reality studies of PM in TBI (e.g., Potvin et al., 2011). Finally, recruiting controls via the personal contacts of the first author may have confounded the results in that there may have been sociodemographic differences between groups. Notwithstanding, effort was devoted to ensuring the sample of controls was representative of the broader community and matched to the TBI group on age, gender and education level.

Despite these limitations, the results of the present research are promising given that virtual platforms offer a range of advantages over their conventional and functional counterparts (e.g., realism and high experimental control), thus are well placed to assess PM deficits in the TBI population. First, virtual reality offers a sensitive and ecologically valid medium through which to assess PM impairment (Lam, Man, Tam, & Weiss, 2006). The VRST is likely sensitive to PM difficulties because it requires sustained cognitive effort over a 14-minute period, incorporates subtasks simulating real world scenarios thereby providing good face validity, places significant demands on attentional processes, and requires the ability to multitask and shift set. Furthermore, examination of the performance of persons with TBI in complex naturalistic environments that are predictive of functional outcomes is more likely to lead to better formulation of the impairments underlying memory failure in individual patients (Kinsella et al., 2009). Second, one of the primary goals of the neuropsychologist in administering tests of PM functioning is to determine whether patients have a normal capacity to remember a simple set of intentions and if not, identify the limits

of their functional independence (Knight et al., 2005). Moreover, neuropsychologists often seek to compare what patients expect to be able to remember against their actual performance. A substantial difference between predicted and actual performance may suggest a lack of self-awareness, which can lead to unrealistic goal-setting during rehabilitation and reduce the motivation of the patient to participate in cognitive retraining. Thus, tests embedded in virtual environments can help patients with disabilities explore the limits of their competence in a non-threatening setting. Third, virtual PM tasks focus on behavioural responses rather than subjective data, thus providing a bridge between conventional neuropsychological tests, behavioural observation and the logistical parameters of in-patient settings (Knight & Titov, 2009). Lastly, virtual platforms and their results have the potential to be readily accepted by test-takers and their families because of their inherent novelty and face validity.

These results have two important implications. First, by exploring the functional-significance of PM impairment; demonstrating the relative difficulty of time- and event-based PM; and exploring the effect of time-monitoring behaviour on time-based PM performance, the current research extends our understanding of the integrity of PM after TBI. Second, these results can be translated into rehabilitative techniques for improving prospective remembering. For example, individuals with TBI could be shown how to convert a time-based task (e.g., taking medication at regular time intervals) into an event-based task (e.g., linking medication to meal times). Although conventional tests of PM are more widely used than virtual platforms, the balance may change as clinicians and researchers become more aware of the benefits of using VR in the assessment and rehabilitation of individuals with TBI. In summary, the results of this study, although preliminary, are novel and promising. The VRST was shown to be a sensitive and ecological measure of time- and event-based PM ability post-TBI.

References

Brandt, J. (1991). The Hopkins Verbal Learning Test: Development of a new memory test with six equivalent forms. *The Clinical Neuropsychologist*, 5, 125-142. doi: 10.1080/138540

49108403297

- Benton, A. L., & Hamsher, K. (1978). *Multilingual Aphasia Examination*. Iowa: Department of Neurology, University of Iowa Hospitals.
- Burgess, P.W. & Shallice, T. (1997). *The Hayling and Brixton Tests*. UK: Thames Valley Test Company.
- Chaytor, N., & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests. *Neuropsychology Review*, 13, 181–197. doi: 10.1023/B:NERV.0000009483.91468.fb
- Craik, F. I. M., & Bialystok, E. (2006). Planning and task management in older adults: Cooking breakfast. *Memory and Cognition*, 34, 1236–1249. doi: 10.3758/BF03193268
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New Jersey: Erlbaum.
- Coltheart, M. (1981). *The MRC Psycholinguistic Database* (Version 2.0) [computer database]. Oxford: Informatics Division Science and Engineering Research Council.
- Elkind, J. S., Rubin, E., Rosenthal, S., Skoff, B., & Prather, P. (2001). A simulated reality scenario compared with the computerized Wisconsin Card Sorting Test: An analysis of Preliminary results. *Cyber Psychology & Behavior*, 4(4), 489–496. doi: 10.1089/109493101750527042.
- Ellis, J. (1996). Prospective memory of the realization of delayed intentions: A conceptual framework for research. In M. A. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), *Prospective memory: Theory and application* (pp. 1-22). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Einstein, G. O., & McDaniel, M. A. (1990). Normal aging and prospective memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 717-726. doi: 10.1037/0278-7393.16.4.717
- Einstein, G. O., & McDaniel, M. A. (2005). Prospective memory: Multiple retrieval processes. *Current Directions in Psychological Science*, 14, 286-290. doi: 10.1111/j.0963-7214.2005.00382.x

- Einstein, G. O., McDaniel, M. A., Richardson, S. L., Guynn, M. J., & Cunfer, A. R. (1995). Aging and prospective memory: Examining the influences of self-initiated retrieval processes. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 21, 996-1007. doi: 10.1037/0278-7393.21.4.996
- Fleming, J. M., Riley, L., Gill, H., Gullo, M. J., Strong, J., & Shum, D. (2008). Predictors of prospective memory in adults with traumatic brain injury. *Journal of the International Neuropsychological Society*, 14, 823-831. doi:10.1017/S1355617708080971
- Goldstein, F. C., & Levin, H. S. (1995). Post-traumatic and anterograde amnesia following closed head injury. In A. D. Baddeley, B. A. Wilson & F.N. Watts (Eds.), *Handbook of memory disorders*. Chichester: John Wiley & Sons.
- Harris, J. E., & Wilkins, A. J. (1982). Remembering to do things: A theoretical framework and an illustrative experiment. *Human Learning*, 1, 12-136.
- Jansari, A., Agnew, R., Akesson, K., & Murphy, L. (2004). Using virtual reality to create an ecologically-valid measure of real-world problems in patients with dysexecutive syndrome. Paper presented at the Symposium to the Joint 27th Conference of the Australian Society for the Study of Brain Impairment and the 27th Mid-Year Meeting of the International Neuropsychological Society Uluru, Australia.
- Kinsella, G. J., Ong, B., & Tucker, J. (2009). Traumatic brain injury and prospective memory in a virtual shopping trip task: Does it matter who generates the prospective memory target? *Brain Impairment*, 10, 45–51. doi:10.1375/brim.10.1.45
- Klinger, E., Chemin, I., Lebreton, S., & Marie, R.-M. (2006). Virtual action planning in Parkinson's Disease: A control study. *Cyber Psychology & Behaviour*, 9(3), 342–347. doi: 10.1089/cpb.2006.9.342.
- Knight, R. G., Harnett, M., & Titov, N. (2005). The effects of traumatic brain injury on the predicted and actual performance of a test of prospective remembering. *Brain Injury*, 19, 27–38. doi:10.1080/02699050410001720022

- Knight, R. G., & Titov, N. (2009). Use of virtual reality tasks to assess prospective memory: Applicability and evidence. *Brain Impairment*, 10, 3-13. doi:10.1375/brim.10.1.3
- Lam, Y. S., Man, D. W. K., Tam, S. F., & Weiss, P. L. (2006). Virtual reality training for stroke rehabilitation. *NeuroRehabilitation*, 21, 245-253.
- Levin, H., & Hanten, G. (2002). Posttraumatic amnesia and residual memory deficit after closed head Injury. In A. D. Baddeley, M. D. Kopelman, B. A. Wilson (Eds.), *The handbook of memory disorders* (pp. 381 - 411). New York: Wiley.
- Logie, R., Trawley, S., & Law, A. (2011). Multitasking: multiple, domain-specific cognitive functions in a virtual environment. *Memory and Cognition*, 39, 1561–1574. doi: 10.3758/s13421-011-0120
- Man, D. W., Fleming, J., Hohaus, L., & Shum, D. (2011). Development of the Brief Assessment of Prospective Memory (BAPM) for use with traumatic brain injury populations. *Neuropsychological Rehabilitation*, 21(6), 884-98. doi: 10.1080/09602011.2011.627270.
- Marsh, R. L., Hicks, J. L., & Bink, M. L. (1998). Activation of completed, uncompleted, and partially completed intentions. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 24, 350-361. doi: 10.1037/0278-7393.24.2.350
- Maujean, A., Shum, D., & McQueen, R. (2003). Effect of cognitive demand on prospective memory in individuals with traumatic brain injury. *Brain Impairment*, 4, 135-145. doi: 10.1375/brim.4.2.135.27024
- Mathias, J. L., & Mansfield, K. M. (2005). Prospective and declarative memory problems following moderate and severe traumatic brain injury. *Brain Injury*, 19, 271-282. doi: 10.1080/02699050400005028
- McCauley, S. R., Pedroza, C., Chapman, S. B., Cook, L., Vásquez, A. C., & Levin, H. (2011). Monetary incentive effects on event-based prospective memory three months after traumatic brain injury in children. *Journal of Clinical Experimental Neuropsychology*, 33(6), 639–646. doi: 10.1080/13803395.2010.547844.

- McDaniel, M. A., & Einstein, G. O. (2007). *Prospective memory: An overview and synthesis of an emerging field*. Thousand Oaks, CA: Sage.
- McGeorge, P., Phillips, L.H., Crawford, J. R., Garden, S. E., Della Sala, S., & Milne, A. B. (2001). Using virtual environments in the assessment of executive dysfunction. *Presence*, 10(4), 375–383. doi: 10.1162/1054746011470235
- National Health and Medical Research Council (NH&MRC). (2007). *National statement on ethical conduct in human research*. Canberra: Australian Government.
- Potvin, M. J., Rouleau, I., Audy, J., Charbonneau, S., & Giguere, J.F. (2010). Ecological prospective memory assessment in patients with traumatic brain injury. *Brain Injury*, 25, 192-205. doi:10.3109/02699052.2010.541896
- Rand, D., Weiss, P., & Katz, N. (2009). Training multitasking in a virtual supermarket: A novel intervention after stroke. *American Journal of Occupational Therapy*, 63, 533-542. doi: 10.5014/ajot.63.5.535
- Rastle, K., Harrington, J., & Coltheart, M. (2002). *358,534 nonwords: The ARC Nonword Database* [computer database]. Australia: Macquarie University.
- Reitan, R. M., & Wolfson, D. (1985). *The Halstead–Reitan Neuropsychological Test Battery: Therapy and Clinical Interpretation*. Tucson: Neuropsychological Press.
- Renison, B., Ponsford, J., Testa, R., Richardson, B., & Brownfield, K. (2012). The ecological and construct validity of a newly developed measure of executive function: The Virtual Library Task. *Journal of the International Neuropsychological Society*, 18, 440-450. doi: 10.1017/S1355617711001883
- Schmitter-Edgecombe, M., Greeley, D. R., & Woo, E. (2009). Characterizing multiple memory deficits and their relation to everyday functioning in individuals with mild cognitive impairment. *Neuropsychology*, 23, 168–177. doi: 10.1037/a0014186
- Schmitter-Edgecombe, M., & Wright, M. J. (2004). Event-based prospective memory following severe closed-head injury. *Neuropsychology*, 18, 353–361. doi: 10.1037/0894-

4105.18.2.353

- Shallice, T., & Burgess, P.W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain*, *114*(2), 727–741.
- Shum, D., Levin, H., & Chan, R. C. (2011). Prospective memory in patients with closed head injury: A review. *Neuropsychologia*, *49*, 2156-2165. doi: 10.1016/j.neuropsychologia.2011.02.006
- Shum, D., Valentine, M., & Cutmore, T. (1999). Performance of individuals with severe long-term traumatic brain injury on time-, event-, and activity-based prospective memory tasks. *Journal of Clinical and Experimental Neuropsychology*, *21*, 49-58. doi: 10.1076/jcen.21.1.49.943
- Smith, R. E. (2003). The cost of remembering to remember in event-based prospective memory: Investigating the capacity demands of delayed intention performance. *Journal of Experimental Psychology*, *29*, 347-361. doi: 10.1037/0278-7393.29.3.347
- Sweeney, S., Kersel, D., Morris, R., Manly, T., & Evans, J. (2010). The sensitivity of a virtual reality task to planning and prospective memory impairments: Group differences and the efficacy of periodic alerts on performance. *Neuropsychological Rehabilitation: An International Journal*, *20*(2), 239-263. doi: 10.1080/09602010903080531
- Tate, R., Hodgkinson, A., Veerabangsa, A., Maggiotto, S. (1999). Measuring psychosocial recovery after traumatic brain injury: psychometric properties of a new scale. *Journal of Head Trauma Rehabilitation*, *14*, 543-557. doi: 10.1097/00001199-199912000-00003
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. New York: Harcourt Brace Jovanovich.
- Wechsler, D. (1997). *Wechsler Memory Scale – III*. United States: The Psychological Corporation.
- Wilson, B., Emslie, H., Foley, J., Shiel, A., Watson, P...Watson, P. (2005). *Cambridge Prospective Memory Test Manual*. Oxford: Harcourt Assessment.
- Woods, S. P., Dawson, M. S., Weber, E., Gibson, S., Grant, I., & Atkinson, J. H. (2009). Timing is

everything: Antiretroviral nonadherence is associated with impairment in time-based prospective memory. *Journal of Clinical and Experimental Neuropsychology*. 15, 42–52.
doi: 10.1017/S1355617708090012

Footnotes

1 In the event that a participant responded to a distractor announcement or responded to an announcement prior to the word ‘sale’ being broadcast the participant was attributed a commission error score of one.

2 Words were generated using the MRC Psycholinguistic Database (Coltheart, 1981), were 4 to 7 characters in length, had a maximum of three syllables, and an age of acquisition rating of 700 (i.e., age 13). Non-words were generated using the ARC Non-word Database (Rastle, Harrington, & Coltheart, 2002), contained 4 to 7 characters, orthographical bodies and polymorphemic syllables.

3 The frequency with which participants checked the time in each minute of the four minutes preceding each time-based task occurring at 4th, 8th and 12th minute of task (i.e., 12 minutes total). No time-based task was present in the last two minutes of the task, thus time checking during this time was not recorded.

Tables

Table 1.

Demographic Details of Participants

Variable	TBI M (<i>Range, SD</i>)	Controls M (<i>Range, SD</i>)
Age	31.68 (18-59, 11.77)	29.72 (21-56, 10.04)
Years of Education	11.71 (9-16, 1.77)	12.52 (10-15, 1.83)
PTA ^a	53.23 (13-251, 50.08)	
GCS	6.74 (3-8, 3.46)	
Time since injury ^a	138 (30-381, 81.94)	

Note. ^aPTA and time since injury were calculated in days.

Table 2.

Performance on PM Tasks and Measures of Neuropsychological Functioning

Measure	TBI M (SD)	Controls M (SD)
<i>LDPMT</i>		
Event-based PM	3.61 (1.63)	5.32 (.69)***
Ongoing component – words	53.13 (6.44)	58.28 (2.21)***
Ongoing component – non-words	51.26 (8.34)	57.64 (2.34)***
<i>VRST</i>		
Time-based PM	2.06 (2.05)	6.36 (1.91)***
Event-based PM	4.84 (2.51)	8.0 (1.68)**
Ongoing component – errands	9.13 (2.90)	11.32 (1.68)***
<i>VRST Time-Checking Behaviour</i>		
1 st Interval	.84 (.78)	1.36 (1.35)
2 nd Interval	1.23 (1.38)	1.56 (1.23)
3 rd Interval	1.16 (1.21)	3.12 (1.54)***
4 th Interval	1.55 (1.67)	4.88 (2.19)***
<i>Cognitive Measures</i>		
Trails A	60.10 (32.88)	24.88 (9.18)***
Trails B	111.85 (49.88)	45.33 (13.39)***
HVLT-R total recall	19.71 (5.66)	26.96 (4.20)***
COWAT	26.06 (10.18)	37.76 (10.66)***
HSCT	4.26 (1.93)	3.28 (.74)***
LNS	9.39 (2.57)	13.80 (3.11)***

Note. ** $p < .01$. *** $p < .001$ for between group comparisons; PM = prospective memory; VRST = Virtual Reality Shopping Task; LDPMT = lexical decision prospective memory test; HVLT-R = Hopkins Verbal Learning Task-Revised; COWAT = Controlled Oral Word Association Test; HSCT = Hayling Sentence Completion Task; LNS = Letter Number Sequencing.

Table 3.

Correlations Between PM Performance and Indices of Neurocognitive and Psychosocial Functioning

	VR TB	VR EB	VR Total	LDPMT EB	VR TB	VR EB	VR Total	LDPMT EB	SPRS OA	SPRS IR	SPRS IL
<i>Controls</i>					<i>TBI</i>						
LDPMT EB	-.312	.584**	.169		.288	.657***	.662***				
Trails A	-.126	-.301	-.294	-.339	.002	-.240	-.198	-.126	.050	.043	.091
Trails B	-.116	.030	-.113	.023	-.190	-.396*	-.412*	-.334	.31*	.019	.37*
HVLT-R-TR	.209	.160	.320	.019	.255	.330	.385*	.480**	.27	.045	.26
COWAT	.058	.002	.012	-.204	.046	.459**	.387*	.483**	.33*	.127	.28*
HSCT	.132	.100	.121	-.020	.055	.309	.274	.224	.039	.195	.110
LNS	-.184	.262	.044	.070	.103	.235	.241	.285	.096	.074	.061
SPRS OA					.602***	-.021	.282	.241			
SPRS IR					-.015	.106	.084	.180			
SPRS IL					.370*	.420*	.512*	.32			

Note. * $p < .05$; ** $p < .01$, *** $p < .001$; VR = Virtual Reality; TB = time-based PM; EB = event-based PM; LDPMT = lexical decision prospective memory test; HVLT-R-TR = Total Recall on the Hopkins Verbal Learning Task-Revised; COWAT = Controlled Oral Word Association Test; HSCT = Hayling Sentence Completion Task; LNS = Letter Number Sequencing; SPRS = Sydney Psychosocial Reintegration Scale (OA = occupational activities; IR = interpersonal relationships; IL = independent living).

Table 4.

Summary of standard regressions of PM performance on the VRST, Trails B and the COWAT on the Occupational Activities and Independent Living Subscales of the SPRS.

Predictor Variable	<i>B</i>	<i>Beta</i>	<i>t</i>	<i>p</i>
<i>Occupational Activities</i>				
VRST TB PM	4.50	.62	3.93	< .001
Trails B	1.43	.43	1.320	.04
COWAT	.33	.27	.320	.66
<i>Independent Living</i>				
VRST TB PM	.99	.41	2.309	.04
VRST EB PM	.33	.23	1.105	.04
Trails B	.01	.15	.758	.28
COWAT	.05	.12	.567	.46

Note. VRST = Virtual Reality Shopping Task; TB = time-based PM; EB = event-based PM; COWAT = Controlled Oral Word Association Test.

Table 5.

Analyses of Variance for User-friendliness Scale

<i>UFS Item</i>	<i>VRST</i>	<i>LDPMT</i>	<i>F</i>	η_p^2	<i>p</i>
	TBI, Controls (<i>M</i> , <i>SD</i>)	TBI, Controls (<i>M</i> , <i>SD</i>)			
<i>Reflectiveness</i>					
PM Task	1.48 (.85), 1.72 (.54)	3.39 (1.45), 4.32 (1.11)	163.47*	.75	< .001
Group			1.68*	.92	< .05
PM Task x Group			3.91	.07	.05
<i>Interesting</i>					
PM Task	1.77 (.96), 1.88 (.97)	3.84 (1.13), 3.64 (.81)	113.56*	.68	<.001
Group			2.06	.001	.81
PM Task x Group			.72	.01	.40
<i>Ease to Understand</i>					
PM Task	1.35 (.61), 1.48 (.59)	1.19 (.48), 1.64 (.57)	.00	.00	.99
Group			6.72*	.11	<.01
PM Task x Group			2.45	.04	.12
<i>Ease to Learn</i>					
PM Task	1.61 (.84), 1.72 (.61)	1.19 (.40), 1.52 (.51)	8.55*	.14	<.01
Group			2.84	.05	.10
PM Task x Group			1.07	.02	.31
<i>Difficulty of Ongoing Task</i>					
PM Task	3.97 (1.22), 4.08 (.91)	3.23 (1.31), 4.12 (.78)	3.26	.06	.08
Group			5.09*	.09	<.05
PM Task x Group			4.05*	.70	<.05
<i>Difficulty of Event Based Task</i>					
PM Task	2.71 (1.24), 4.52 (.71)	3.39 (.95), 4.28 (.74)	5.11*	.09	<.05
Group			43.44*	.19	<.001
PM Task x Group			12.45*	.45	<.001
<i>Recommendation</i>					
PM Task	2.13 (1.38), 1.56 (.65)	3.84 (1.16), 4.08 (1.04)	84.28*	.61	<.001
Group			.74	.01	.39
PM Task x Group			3.09	.05	.08

Note. $Df = 1, 54$ for all analyses. Items were rated on a 5-point Likert Scale whereby '1' denotes strongly agree and '5' denotes strongly disagree. VRST = Virtual Reality Shopping Task. LDPMT = Lexical Decision Prospective Memory Task.

Figures

Figure 1. Screen shot of the inside of the VR shopping centre. The floor plan is located in the upper left corner, task list in the upper right corner, and mobile phone in the bottom right corner

Figure 2. Screen shot of the virtual mobile phone. Text messages are sent by clicking on the green “SEND” button.

Appendix A

Evaluation of Computer Tasks

Thank you for participating in this study. We are interested in your feedback on the two computer tasks that you have just done. Please respond by circling your response for the following questions.

Shopping Task

- 1 The task included activities that were similar to what I would do in my everyday life.**

Strongly Agree

Strongly Disagree

1

2

3

4

5

- 2 I found the task interesting.**

Strongly Agree

Strongly Disagree

1

2

3

4

5

- 3 The instructions for the task were clear and easy to understand.**

Strongly Agree

Strongly Disagree

1

2

3

4

5

- 4 The task was easy to learn.**

Strongly Agree

Strongly Disagree

1

2

3

4

5

- 5 I found the shopping part of the task difficult.**

Strongly Agree

Strongly Disagree

1

2

3

4

5

- 6 I found the sales announcement part of the task difficult.**

Strongly Agree

Strongly Disagree

1

2

3

4

5

- 7 I found the mobile phone messaging part of the task difficult.**

Strongly Agree

Strongly Disagree

1

2

3

4

5

8 I would recommend this task to other people.

Strongly Agree

Strongly Disagree

1

2

3

4

5

Word Decision Task**1 The task included activities that were similar to what I would do in my everyday life.**

Strongly Agree

Strongly Disagree

1

2

3

4

5

2 I found the task interesting.

Strongly Agree

Strongly Disagree

1

2

3

4

5

3 The instructions for the task were clear and easy to understand.

Strongly Agree

Strongly Disagree

1

2

3

4

5

4 The task was easy to learn.

Strongly Agree

Strongly Disagree

1

2

3

4

5

5 I found the word decision part of the task difficult.

Strongly Agree

Strongly Disagree

1

2

3

4

5

6 I found the animal word part of the task difficult.

Strongly Agree

Strongly Disagree

1

2

3

4

5

7 I would recommend this task to other people.

Strongly Agree

Strongly Disagree

1

2

3

4

5