Validation of the virtual-reality prospective memory test (Hong Kong Chinese version) for
individuals with first-episode schizophrenia

Author Note

(Paper re-submitted to *Neuropsychological Rehabilitation* for review and possible publication.)

David W. K. Man¹, Balasankar Ganesan², Calvin C. K. Yip¹, Christina O. P. Lee³, Sarah Y. L. Tsang⁴, Pan W. P. Yu⁵, Janice G. E. Young⁶, David H. K. Shum⁷

¹ Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hung Hom, Hong Kong

² Center of Health Technology, FEIT, University of Technology Sydney, Australia

³ Hope Community Rehabilitation Day Centre, Haven of Hope Christian Service, Hong Kong

⁴ Placidity Place (ICCMW), The Society of Rehabilitation and Crime Prevention, Hong Kong.

⁵ The Tsung Tsin Mission of Hong Kong Social Service, Hong Kong

⁶ Mercy Medirehab Professional Group Limited

⁷ Behavioural Basis of Health Program, Griffith Health Institute, Griffith University, Gold Coast, Australia

* Correspondence should be addressed to: Prof David W. K. Man. Postal address: Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Yuk Choi Road, Hung Hom, Kowloon, Hong Kong. Telephone: (852) 276-66711. Fax: (852) 2330-8656. Email address: David.Man@polyu.edu.hk.

Abstract

This study was performed to examine the psychometric properties of a Virtual-Reality Prospective Memory Test (Hong Kong Chinese version; VRPMT-CV). The VRPMT was administered to 44 individuals with first-episode schizophrenia. The test was administered again 2 weeks later to establish test-retest reliability. The concurrent validity of the VRPMT was evaluated by examining the correlations between the VRPMT score and the score on the Chinese version of the Cambridge Prospective Memory Test (CAMPROMPT-CV). The performance of individuals with schizophrenia on the VRPMT was also compared with that of 42 healthy control subjects to examine the test's sensitivity and specificity. The intra-class correlation for test-retest reliability of the total VRPMT-CV score was 0.78 ($p = 0.005$). A significant correlation was found between the total VRPMT-CV score and the total CAMPROMPT-CV score ($r = 0.90$; $p < 0.001$). Comparison with the healthy control subjects revealed that the total VRPMT-CV score was a sensitive (92.9%) and specific (75%) measure of prospective memory deficits in individuals with schizophrenia. The VRPMT-CV is an assessment of prospective memory that has good construct validity, test-retest reliability, sensitivity and specificity in the context of first-episode schizophrenia.

Keywords: First-onset Schizophrenia, Prospective Memory, Virtual Reality Memory Test, Reliability, Validity, Sensitivity, Specificity

Validation of the virtual-reality prospective memory test (Hong Kong Chinese version) for individuals with first-episode schizophrenia

Prospective memory (PM) deficits are amongst the possible cognitive impairments in people with schizophrenia (Au et al., 2013). PM is the memory required to perform an intended action at some point in the future, in response to an external event or at a specific time (Radford, Lah, Say & Miller, 2011). PM tasks are classified into event-based (e.g., buying a book when passing by a bookshop), time-based (e.g., attending an appointment at 11 AM) and activity-based (e.g., taking medication after finishing a meal) tasks (Ellis, 1996). PM deficits can lead to poor medication adherence (Moore, Sellwood & Stirling, 2000) and can affect scores on the Activities of Daily Living test and the Instrumental Activities of Daily Living test as well as work and social life (Au et al., 2014; Ordeman, Opper & Davalos, 2014). Early identification of PM problems and tailor-made interventions are therefore considered crucial to the rehabilitation of patients with schizophrenia.

PM impairments have been assessed with self-report questionnaires or psychometric tests. The Prospective and Retrospective Memory Questionnaire (Hsu & Hua, 2011; Smith, Della Sala, Logie & Mayor, 2000) and the Comprehensive Assessment of Prospective Memory Questionnaire (Chan, Qing, Wu & Shum, 2010) are examples of this type of questionnaire, but they are not suitable for individuals with schizophrenia who might have limited insight into their condition (Chan et al., 2008). Objective paper-and-pencil tests such

as the Rivermead Behavioural Memory Test (Wilson, Cockburn & Baddeley, 2003), the Cambridge Prospective Memory Test (CAMPROMPT; Wilson, Emslie, Foley, Shiel & Watson, 2005); the Memory for Intentions Screening Test (Raskin, 2004) and the Royal Prince Alfred Prospective Memory Test (Radford, Lah, Say & Miller, 2011) have been the focus of research.

However, a limitation of these assessments is that they may have little predictive value for patients' performance in a real-life situation. Thus, ecological validity is considered crucial in PM assessment. Ecological validity in neuropsychological assessment can be conceived in terms of two factors, verisimilitude and veridicality (Chaytor & Schmitter-Edgecombe, 2003). Verisimilitude is the degree of resemblance between a cognitive task and its demands and the everyday environment, whereas veridicality is the extent to which task performance predicts everyday functioning. One technique that can address better ecological validity is virtual reality (VR). Virtual environments (VEs) have been suggested as potential aids to enhance the ecological validity of such assessments (Parson, 2015). Their sensitivities in the detection of cognitive impairments have been supported (Negut, Matu, Sava & David, 2016). Several ecologically valid measures have been documented (Parsons, Carlew, Magtoto & Stonecipher, 2015), including the Jansari Assessment of Executive Function (Jansari et al., 2014). VR measures allow for real-time assessment of cognitive processing in a manner that can more closely resemble real-life activities (Matheis et al., 2007). The key

factor is related to the concept of presence (Parsons et al., 2015). VEs allow the users to experience virtual situations that appear to be real-life while being assessed in a protected, controlled and safe context. The use of VR has been studied in the assessment and clinical practice of psychoses (Kim et al., 2010; Valmaggia, Day & Rus-Calafell, 2016; Veling, Pot-Kolder, Counotte, Os & Gaag, 2016). Specific measurement of memory and executive function in people with schizophrenia has also been documented (Ku et al., 2004; Siemerks et al., 2012). They have been found to minimise the social desirability bias. The attractive user interface also reduces the users' anxiety levels, improves motivation and is helpful for assessment of people with poor verbal communication skills (Kim et al., 2010). Examples of recently developed VR-based PM assessment include the Virtual Library Task (Renison, Ponsford, Testa, Richardson & Brownfield, 2012); the Test Ecologique de Mémoire Prospective (Potvin, Rouleau, Audy, Charbonneau & Giguère, 2010), the Virtual Bungalow (Attree et al., 2009), the Virtual Week (Rendell & Craik, 2000) and the Virtual Reality Prospective Memory Test (VRPMT), which was evaluated in this study. We aimed to validate the utility and suitability of the VRPMT (Hong Kong Chinese version; VRPMT-CV) for use with individuals with schizophrenia. Specifically, we examined the psychometric properties of the VRPMT-CV, namely, its concurrent validity, test-retest reliability and sensitivity and specificity to PM deficits in schizophrenia. We also investigated differences between performance on event-based and time-based PM tasks, within and between the first-episode

schizophrenia group and the healthy control group. The research hypotheses are as follows:

1. Concurrent validity. There are statistically significant differences in correlation coefficients (r) between the PM performance of individuals with first-onset schizophrenia as measured by the VRPMT-CV and the Chinese version of the CAMPrompt (CAMPrompt-CV).
2. Test-retest reliability. There are no statistically significant differences between the mean PM performance in the initial testing and retesting of the VRPMT-CV.
3. Sensitivity and specificity to PM deficits in schizophrenia. There are statistically significant differences in the PM performance between individuals with first-onset schizophrenia and those in the healthy control groups.
4. There are statistically significant differences between event-based and time-based PM performance within and between the patients in the schizophrenia group and those in the healthy control group.

Methods

This study of the VRPMT-CV was divided into three parts: (1) translation and evaluation of the Hong Kong Chinese version, (2) evaluation of its concurrent validity and reliability and (3) evaluation of its sensitivity and specificity.

Part 1: Linguistic and Cultural Validity of the Translated Version

Participants. The review panel of three members included one psychologist and two occupational therapists who had experience working in psychiatric rehabilitation for more than 5 years.

Measures. The panel members independently rated the equivalence, clarity and relevance of 102 translated test items of the VRPMT-CV on a six-point Likert scale (1 = strongly disagree, 2 = disagree, 3 = slightly disagree, 4 = slightly agree, 5 = agree, 6 = strongly agree).

Procedures. The VRPMT-CV was first accurately translated and appropriately adapted from the English language original. Written instructions and auditory scripts were translated from English into Hong Kong Chinese (Cantonese) with adaptations according to local culture, such as replacing references to the film star Tom Cruise in two of the event-based PM cues with references to Jacky Chan, a Hong Kong film star, and changing the format of a car number plate from the Australian to the Hong Kong format. The translated version was then evaluated by the expert panel.

Part 2: Validity and Reliability Study

Participants. Convenience sampling was used to recruit 44 participants with first-episode schizophrenia from three non-governmental organisations: Halfway House III and

Joyous Place Hostel of the New Life Psychiatric Rehabilitation Association; Yeung Sing Memorial Long Stay Care Home of Tung Wah Group of Hospitals; and Caritas Jockey Club Lai King Rehabilitation Centre. The inclusion criteria were age between 20 and 50 years and an understanding of spoken and written Cantonese, the Chinese dialect of Hong Kong. All participants had received a diagnosis of first-episode schizophrenia from a psychiatrist using the DSM-IV. Patients were required to be clinically stable: they should not have been in a psychiatric hospital in the past 6 months, and their antipsychotic medication should have been stable for the past 6 months. Also, they should not suffer from acute depression at the time of the assessment. It had been less than 5 years from onset in all patients, and all patients had received second-generation antipsychotics. The exclusion criteria were mild cognitive impairment or early dementia, significant medical conditions, epilepsy, auditory or visual impairments, a history of alcohol or substance abuse, and electroconvulsive therapy within the past year.

Procedures. Approval was given by the Ethics Committee of the Hong Kong Polytechnic University. All participants provided written consent before the assessments were conducted and were offered an incentive of HK\$30 per session. After screening for general cognitive impairment with the Mini Mental State Examination – Hong Kong Chinese version, the participants completed the VRPMT-CV and the CAMPRMPT-CV; the order in which the tests were administered was counterbalanced. All participants were given clear

verbal instructions before both tests. They were allowed to complete a maximum of two computer tutorials before the assessment phase of the VRPMT-CV to ensure that they had a good understanding of all task requirements. The participants were given a 5- to 10-minute break between the two tests to minimise the effects of mental fatigue. All participants were retested on the VRPMT-CV within 2 weeks.

Measures. The following measures were used.

Mini Mental State Examination – Hong Kong Chinese version (Chiu et al., 1994).

This test was used as a screening tool to exclude patients whose global cognitive functioning was too low to enable them to participate in the study. Patients with a score of 20 or lower were excluded.

Chinese version of CAMPROMPT (CAMPROMT-CV; Man et al., 2015; Wilson et al., 2005). This is one of the few Chinese PM tests available; it was used as a reference point to establish the concurrent validity of the VRPMT-CV. It consists of three event-based and three time-based PM tasks; paper-and-pencil quizzes and puzzles are used as ongoing distraction tasks, as in the original English version. All time-based or event-based tasks were scored from 0 to 6 on the basis of eight scoring criteria; the range of total scores for the six tasks was 0 to 36, with higher scores indicating better PM performance. The test took 25 minutes to administer. It has been reported to have good reliability and validity (Wilson et al., 2005) and has been used to assess PM function in individuals with schizophrenia (Au et al., 2014).

Virtual Reality Prospective Memory Test – Chinese version (VRPMT-CV). This test was adapted from the original English version called the Virtual Reality Shopping Test (VRST), which was developed by a team including the first and last authors of this study and has been previously used to assess PM in individuals with traumatic brain injury (Canty et al., 2014). The VRST has sound convergent validity; it demonstrated a positive relationship between PM performance and standardised cognitive tests. Its ecological validity was shown by its being more reflective of everyday activities performance. It can be a sensitive measure of PM dysfunction following traumatic brain injury. Its VRST time-based PM cut-off score of 5 or lower indicated a sensitivity of 76%, a specificity of 71%, a positive predictive value of 88% and a negative predictive value of 65% (Canty et al., 2014).

The VRPMT-CV was developed from the VRST for use in the unique linguistic and cultural environment of Hong Kong. This 14-minute test is based on a shopping centre scenario and comprises three time-based and three event-based PM tasks. The distraction task is to navigate 20 different shops to obtain items from the shopping list. The digital clock on the mobile phone was always set to 8:00 PM at the start of the test. The time-based PM tasks required the participants to send text messages 4, 8 and 12 minutes after the start of the test, using the digital clock on the mobile phone to keep track of time. Each correct response within 10 seconds was scored 1, otherwise no score was given. A maximum score of 3 would be allocated for the time-based PM task. Figure 1 shows how the time-based PM task was

executed. The event-based PM tasks required the participants to press the 'T' key in response to three sales announcements that were made during the 3rd, 9th and 13th minutes; three other non-sales announcements were broadcast during the 5th, 7th and 11th minutes as distracters. A correct response to each of the announcements was awarded one point (maximum event-based PM score = 3). Figure 1b shows how the event-based PM task was executed. Before the test phase, the participants completed a short computer tutorial to familiarise themselves with the environment and the tasks. A detailed performance summary is generated automatically by the VRPMT programme at the end of a test; the summary includes the scores on the time-based and event-based PM tasks, the number of items purchased and the frequency with which the participant checked the time.

The VRPMT-CV is a non-immersive VR programme that was developed using 3DVIA Virtools game development software. The VR includes background noise from a shopping centre to increase the verisimilitude. The participants navigate the shopping centre using the mouse and three clickable icons (shopping centre map, shopping list and mobile phone) located at three corners of the screen.

Figure 1 about here

Part 3. Sensitivity and Specificity Study

Participants. In addition to the participants with schizophrenia recruited for Part 2, another group of 42 healthy participants was recruited and completed the same tests in a similar time frame. They were normal, healthy adults recruited from the general population using convenience sampling via advertisement. Attempts were made to match the two groups in terms of age, gender and education level. The same inclusion and exclusion criteria were adopted. Additional exclusion criteria for the healthy control group included a diagnosis of mental illness and a first-degree relative with a history of schizophrenia or schizoaffective disorder.

Measures. The same instruments were used as in Part 2.

Procedures. The same procedures were applied to the health control group as in Part 2.

Data analysis. As a preliminary step, all data were screened to determine whether they met the criteria for parametric analysis, and appropriate nonparametric tests were used when necessary. Two-tailed tests with a significance level (p value) of less than 0.05 were used in all analyses.

The intraclass correlation coefficient was used to evaluate the quality of the translation, and the limits of agreement were used to evaluate the test-retest reliability. A two-way random-effects model of intraclass correlation coefficient, which took into account both the

correlation of the scores and the agreement of the scores, was used to examine the reliability of the expert panel members in rating three areas (equivalence, clarity and relevancy) of all translated items. The items were independently scored by four expert panel members. Similarly, a two-way mixed-effects model of intraclass correlation coefficient was used to determine the test-retest reliability of the VRPMT-CV.

The concurrent validity of the VRPMT-CV in the schizophrenia group was assessed by calculating Spearman's rho value for correlations between the time-based, event-based and total scores on the VRPMT-CV and the CAMPROPT-CV. The time-based, event-based and total scores on the VRPMT-CV were obtained at two time points (first test and retest) from the 44 participants in the schizophrenia group. The performance of the schizophrenia and healthy control groups on the two types of PM tasks was compared using a two (group: control, schizophrenia) by two (task type: time-based PM, event-based PM) analysis of variance with demographic information used as covariates. If these variables did not differ statistically in the two groups, they were not included in the analysis of covariance. The sensitivity and specificity of the VRPMT-CV to PM deficits in the schizophrenia group were evaluated with receiver operating characteristic (ROC) curve analysis. The area under the curve (AUC) was calculated to determine the diagnostic value of the ROC curve. Mann-Whitney U tests were used to assess group differences in the number of time checks and the number of shopping tasks completed. Moreover, a Mann-Whitney U test was used to

examine differences in computer-generated data between the two participant groups in the number of time checks and the number of shopping tasks completed.

Results

Quality of Translation

The translated items were divided into two main categories: those related to the six PM tasks (three time-based and three event-based) and those related to the distraction task and the VE (e.g., announcements, signs, names of shops and pop-up windows showing messages and response options).

Table 1 shows the agreement between the expert panel members regarding the translated material. Moderate agreement was found in three aspects of the translation, suggesting an acceptable rating of the quality of the translation (i.e., the extent to which it is equivalent to the English version, has a clear meaning and has cultural relevance).

Table 1 about here

Group Comparison on PM Performance

Table 2 summarises the demographic characteristics and PM performance of the schizophrenia and the healthy control groups. Group comparisons using the Mann-Whitney U test and chi-square test of independence showed no significant group differences in age or gender, but a significant difference was seen in the education level ($MWU = 225$; $z = -3.04$; $p = 0.02$) due to our limited control over the characteristics of the volunteer participants in the healthy control group.

Table 2 about here

PM was divided into event-based and time-based components. Analysis of variance using the group as the between-subjects factor and the task type as the within-subject factor indicated similar performance on the two types of PM task [$F(1,52) = 0.08$; $p = 0.09$], and no group by task type interaction was seen [$F(1,52) = 0.92$; $p = 0.34$]. However, between-subjects testing showed significant differences between the schizophrenia and healthy control groups in their event-based and time-based PM performance [$F(1,52) = 33.52$; $p < 0.001$].

The level of education [$F(1,52) = 0.92$; $p = 0.34$] did not significantly contribute to the differences in performance in event-based and time-based PM. Further sub-group analysis indicated no statistically significant differences between time-based and event-based PM

performance for either group ($p = 0.12$ for the schizophrenia group; $p = 0.06$ for the healthy control group).

The number of shopping tasks completed in the schizophrenia group (mean, 4.46; $p < 0.001$) was significantly lower than in the healthy group (Mann-Whitney U test; mean, 9.93; $p < 0.001$). Significantly fewer time-checking events were seen in the schizophrenia group (mean, 4.89; $p < 0.001$) than in the healthy control group (mean, 9.86; $p < 0.001$).

Figure 2 about here

Concurrent Validity

Table 3 shows that the VRPMT-CV scores in both the schizophrenia group and the healthy control group were positively correlated with the CAMPrompt-CV scores.

Table 3 about here

Test-Retest Reliability

Table 1 shows that the schizophrenia group's VRPMT-CV initial test and re-test performances were positively correlated with respect to the time-based tasks, event-based tasks and total performance (sum of scores on both time-based and event-based tasks). These

results indicate that the VRPMT-CV has good test-retest reliability with respect to time-based and event-based PM tasks in this sample of patients. The test-retest reliability of the VRPMT-CV scores also showed good reliability between the two time points in the healthy control group.

Sensitivity and Specificity

In the schizophrenia group, 64.3% had CAMPROMPT-CV scores below the cutoff, indicating that they suffered from PM problems. However, in the healthy control group, only 7.1% suffered from PM problems. The ROC curve of the VRPMT-CV was analysed with the cutoff scores of the CAMPROMPT-CV as a gold standard for PM problems.

Using a cutoff score of 3.5 for the VRPMT-CV, the sensitivity and specificity to PM problems were found to be 83.3% and 85%, respectively, and the AUC was 0.944 ($p < 0.001$). Using the VRPMT-CV event-based score of 1.5, the sensitivity and specificity were 94.4% and 70%, respectively, and the AUC was 0.901 ($p < 0.001$). For a score of 1.5 on the VRPMT-CV time-based PM, the sensitivity and specificity were 83.3% and 80%, respectively, and the AUC was 0.854 ($p < 0.001$). To increase the specificity of the time-based and event-based PM, the cutoff points of the VRPMT-CV event-based score and time-based score were increased. When we adopted a cutoff of 2.5 for event-based PM, the sensitivity and specificity were changed to 75% and 90% respectively. For the time-based

score, an increase in the cutoff to 2.5 changed the sensitivity and specificity to 52.8% and 95%, respectively. Therefore, the time-based and event-based scores of the VRPMT-CV could have higher specificity as a trade-off for a decrease in sensitivity (Figure 3).

Figure 3 about here

Amongst the three AUCs of the ROC curves, the AUC of the total VRPMT-CV score had the largest area, which indicates that the most informative and useful diagnostic value could be obtained from the total score. It is thus suggested that the total score might be better than the time-based or event-based scores to screen for PM impairment in individuals with schizophrenia.

DISCUSSION

Our preliminary findings suggest that the VRPMT has good convergent validity and is a sensitive measure of PM dysfunction in people with schizophrenia. The hypothesis that performance on the VRPMT-CV would correlate with performance on the CAMPRMPT-CV in patients with schizophrenia was supported. The use of a VE has clinical utility that may be comparable to conventional PM measures and possibly better than is typically attributed to VR. Hypothesis two, that people with schizophrenia would perform similarly on the VRPMT-CV in the first test and re-test, was partially supported. An additional finding

was that the total score on the time-based tasks showed slightly better test-retest reliability in people with schizophrenia.

Hypothesis three, that significant differences would be seen in PM performance between individuals with first-onset schizophrenia and the healthy control group, was confirmed. The sensitivity and specificity of the VRPMT-CV were also supported.

Finally, the study found preliminary evidence of differences in PM measures between the patients and the healthy control group. Individuals with schizophrenia performed less well on the VRPM-CV than those in the healthy control group. This finding is consistent with previous research and suggests that PM impairment is a significant cognitive problem in patients with schizophrenia and affects their everyday functioning. It was found that PM impairment mainly originated in problems with cue detection and intention retrieval (Wang et al., 2008; Woods et al., 2007). These problems might be present in people with schizophrenia due to their working memory impairment, but further investigation should be conducted because the intraclass correlation coefficient alone was not sufficient to support this assumption. It has also been argued that cue detection and intention retrieval depend upon the ability to self-initiate actions; the prefrontal lobe is involved in self-initiation, and prefrontal abnormalities have been reported in patients with schizophrenia (Wang et al., 2008). We did not observe great differences in performance on time-based and event-based PM tasks, possibly because there were too few test items or items of a wider level of difficulty. The

narrow range of the subscores (0 to 3) could be another possible reason. However, it should be noted that there was a trend towards better performance on the event-based tasks, suggesting that the time-based tasks were a bit more difficult (Ordeman, Opper & Davalos, 2014). Perhaps time-based tasks are more abstract and less dependent on retrospective memory, and successful performance cannot be based on external cues (Groot et al., 2002; Shum et al., 2004; Wang et al., 2008). Moreover, it has been shown that individuals with schizophrenia have time processing deficits, which may have a greater effect on time-based PM than event-based PM (Wang et al., 2008). Sensitivity analysis suggests that PM impairment in participants with schizophrenia is greater than that in healthy controls when the number of correct responses on event-based and time-based tasks in the VRPMT-CV is compared. Thus, management of PM impairment in people with schizophrenia can be more specifically geared towards poorer time-based prospective remembering.

Compared to pencil-and-pen PM tests or self-report type questionnaires, the VRPMT-CV has several positive features. It is a convenient tool that can be used on any standard personal computer, notebook or tablet PC. The result of an assessment is automatically generated by the programme that runs the test, so scoring requires minimal time and effort and is accurate, reliable and objective. The VRPMT-CV is considered more ecologically valid than traditional paper-and-pencil PM assessments because the VE is a familiar one and the PM tasks are part of most participants' everyday experience. Another study has shown that the results of VR

assessments correlate significantly with everyday functioning (Canty et al., 2014). However, our data do not support the previous findings that VRPMT-CV is associated with daily functioning because this association was not assessed in this study. Further study of the relationship of the VRPMT-CV with functional independence and psychosocial functioning in addition to cognitive measures is recommended.

The VRPMT-CV has an important advantage in that it is easier for people with poor communication or social skills, such as people with first-episode schizophrenia, to complete. Their stress and anxiety associated with communicating with the assessor might be greatly reduced, and thus more accurate PM results can be obtained. Therefore, it is suggested that the VRPMT-CV could be administered to people with schizophrenia, apart from the population with traumatic brain injury, by using the VRST (the original version of the VRPMT-CV).

Our study has some limitations. First, we did not record detailed information about patient medication such as chlorpromazine-equivalent dosages of antipsychotic medication or the types of medication the patients were receiving at the time of this study. This may give rise to questions about whether PM performance was influenced by the side-effects of medication. However, a recent review of PM in schizophrenia suggested that PM deficits are not related to medications associated with schizophrenia (Ordeman et al., 2014). Second, we did not test mood conditions, such as with the positive and negative syndrome scale, although

we have operationally defined a stable mood condition. The absence of positive and negative syndrome scale scores (or scores from another equivalent instrument) does not enable comparison with other samples. Third, even after completing two tutorial sessions, a few of the participants in the schizophrenia group required additional explanation in the use of the VRPMT-CV, and variations in the assessors' additional explanations may have influenced the performance in these cases. Fourth, there is no alternate version of the VRPMT-CV to minimise practice effects when reassessment is required. This becomes a shortcoming with respect to the assessment of event-based PM.

We recommend a number of modifications to the VRPMT-CV. First, the computer tutorial should include both auditory and written instructions to make it easier for participants to understand the test procedure and task requirements. This could reduce the role of assessors' supplementary instructions, which are a potential source of variance in performance. Second, the development of alternative versions using different distraction tasks might improve the usability of the test.

CONCLUSIONS

This study shows that the VRPMT-CV has good psychometric properties despite its methodological limitations. The VRPMT-CV has the potential to be a valid and reliable tool to screen or assess for PM function for schizophrenia patients in Hong Kong and other

Chinese-speaking patients with schizophrenia. Further study in developing norms for schizophrenia and other psychiatric groups, such as bipolar disorders and depression, is recommended. This would help to improve the rate of problem detection and of decision making in clinical neuropsychological rehabilitation.

ACKNOWLEDGEMENT

This study was partly supported by a Research Grant awarded by the Department of Psychology, Griffith University, Australia, to the first author. The University had no further role in the design of the study, in the collection and interpretation of data, in writing the report or in the decision to submit the paper for publication.

The authors thank the staff members of Halfway House III and Joyous Place Hostel of New Life Psychiatric Rehabilitation Association, TWGHs Yeung Sing Memorial Long Stay Care Home and Caritas Jockey Club Lai King Rehabilitation Centre for providing staunch support in our research project.

REFERENCES

Attree, E. A., Dancey, C. P., & Pope, A. L. (2009). An assessment of prospective memory retrieval in women with chronic fatigue syndrome using a virtual-reality environment: an initial study. *CyberPsychology and Behavior* 12(4), 379–385.

Au, R. W. C., Man, D., Shum, D., Lee, E., Xiang, Y. T., Ungvari, G. S., & Tang, W. K. (2014). Assessment of prospective memory in schizophrenia using the Chinese version of the Cambridge Prospective Memory Test: A controlled study. *Asia-Pacific Psychiatry* 6, 54–61.

Au, R. C. W., Ungvari, G. S., Lee, E., Man, D., Shum, D., Xiang, Y. T., & Tang, W. S. (2013). Prospective memory impairment and its implications for community living skills in bipolar disorders. *Bipolar Disorder* 15(8), 885–892.

Canty, A. L., Fleming, J., Patterson, F., Green, H., Man, D., & Shum, D. (2014). Evaluation of a virtual reality prospective memory task for use with individuals with severe traumatic brain injury. *Neuropsychological Rehabilitation* 24(2), 238–265.

Chan, R. C. K., Qing, Y. H., Wu, Q. P., & Shum, D. (2010). Prospective memory in healthy Chinese people: The latent structure of the Comprehensive Assessment of Prospective Memory Questionnaire. *Neuropsychological Rehabilitation* 20(3), 459–470.

Chan, R. C. K., Wang, Y., Ma, Z., Hong, X. H., Yuan, Y., Yu, X., Li, Z., Shum, D., & Gong, Q. Y. (2008). Objective measures of prospective memory do not correlate with subjective complaints in schizophrenia. *Schizophrenia Research* 103(1–3), 229–239.

Chaytor, N., & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests. *Neuropsychology Review* 13, 181–197.

- Chiu, H. F. K., Lee, H. C., Chung, W. S., & Kwong, P. K. (1994). Reliability and validity of the Cantonese version of mini-mental state examination: a preliminary study. *Hong Kong Journal of Psychiatry* 4(2), 25–28.
- Ellis, J. (1996). Prospective memory or the realization of delayed intentions: A conceptual framework for research. In M. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications* (pp. 1–22). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Groot, Y., Wilson, B., Evans, J., & Watson, P. (2002). Prospective memory functioning in people with and without brain injury. *Journal of the International Neuropsychological Society* 8, 645–654.
- Hsu, Y. H., & Hua, M. S. (2011). Taiwan version of the Prospective and Retrospective Memory Questionnaire: Latent structure and normative data. *Clinical Neuropsychology* 26, 240–249.
- Jansari, A. S., Devlin, A., Agnew, R., Akesson, K., Murphy, L., & Leadbetter, T. (2014). Ecological assessment of executive functions: A new virtual reality paradigm. *Brain Impairment* 15(2), 71–87.
- Kim, K., Kim, S. I., Cha, K. R., Park, J., Rosenthal, M. Z., Kim, J. J., Han, K., Kim, I. Y., & Kim, C. H. (2010). Development of a computer-based behavioral assessment of checking behavior in obsessive-compulsive disorder. *Comprehensive Psychiatry* 51, 86–93.
- Ku, J., Cho, W. G., Kim, J. H., Kim, K. U., Kim, B. N., Hahn, W. Y., Kim, J. J., Lee, J. H., Kim, I. Y., & Kim S. I. (2004). The development of a VR system for the cognitive and behavioral assessment of schizophrenia. In J. D. Westwood et al. (Eds.), *Studies in health technology and informatics* (pp. 180–182). Amsterdam, Netherlands: IOS Press.

VALIDATION OF HKC VR PROSPECTIVE MEMORY TEST FOR SCHIZOPHRENIA 26

Man, D. W. K., Yip, C. C. K., Lee, Y. Y. G., Shum, D. (2015). Validation of the Cambridge Prospective Memory Test (Hong Kong Chinese version) for people with stroke. *Neuropsychological Rehabilitation* 25(6), 895–912.

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

Moore, A., Sellwood, W., & Stirling, J. (2000). Compliance and psychological resistance in schizophrenia. *British Journal of Clinical Psychology* 39, 287–295.

Negut, A., Matu, S. A., Sava, F. A., & David, D. (2016). Virtual reality measures in neuropsychological assessment: A meta-analysis review. *The Clinical Neuropsychologist* 39(2), 165–184.

Ordeman, G. J., Oppen, J., & Davalos, D. (2014). Prospective memory in schizophrenia: A review. *Schizophrenia Research* 155, 77–89.

Parson, T. D. (2015). Virtual reality for enhanced validity and experimental control in the clinical, affective and social neurosciences. *Frontiers in Human Neuroscience* 9, 1–19.

Parson, T. D., Carlew, A. R., Magtoto, J. & Stonecipher, K. (2015). The potential of function-led virtual environments for ecologically valid measures of executive function in experimental and clinical neuropsychology. *Neuropsychological Rehabilitation*, 11, 1-31. doi: 10.1080/09602011.2015.1109524

Potvin, M. J., Rouleau, I., Audy, J., Charbonneau, S., & Giguère, J. F. (2010). Ecological prospective memory assessment in patients with traumatic brain injury. *Brain Injury* 25, 192–205.

- Radford, K. A., Lah, S., Say, M. J., & Miller, L. A. (2011). Validation of a new measure of prospective memory: the Royal Prince Alfred Prospective Memory test. *Clinical Neuropsychologist* 25(1), 127–140.
- Raskin, S. (2004). Memory for intentions screening test. *Journal of the International Neuropsychological Society* 10(Suppl. 1), 110.
- Rendell, P. G., & Craik, F. I. M. (2000). Virtual week and actual week: Age-related differences in prospective memory. *Applied Cognitive Psychology* 14(7), S43–S62.
- 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.
- Shum, D., Ungvari, G. S., Tang, W. K., & Leung, J. P. (2004). Performance of schizophrenia patients on time-, event-, and activity-based prospective memory tasks. *Schizophrenia Bulletin* 30, 693–701.
- Siemerkus, J., Irle, E., Schmidt-Samoa, C., Dechent, P., & Weniger, G. (2012). Egocentric spatial learning in schizophrenia investigated with functional magnetic resonance imaging. *NeuroImage: Clinical* 1, 153–163.
- Smith, G., Della Sala S., Logie, R. & Mayor, E.A. (2000). Prospective and retrospective memory in normal aging and dementia: a questionnaire study. *Memory* 8, 311–321.
- Valmaggia, L. R., Day, F., & Rus-Calafell, M. (2016). Using virtual reality to investigate psychological processes and mechanisms associated with the onset and maintenance of psychosis: A systematic review. *Social Psychiatry and Psychiatric Epidemiology*, 51, 921-936. doi: 10.1007/s00127-016-1245-0.

VALIDATION OF HKC VR PROSPECTIVE MEMORY TEST FOR SCHIZOPHRENIA 28

1
2
3 Veling, W., Pot-Kolder, R., Counotte, J., Os, J. V., & van der Gagg, M. (2016). Environmental
4
5 social stress, paranoia and psychosis liability: a virtual reality study. *Schizophrenia*
6
7 *Bulletin*, 42, 1363-1371. doi: 10.1093/schbul/sbw031
8
9
10 Wang, Y., Chan, R. C., Hong, X., Ma, Z., Yang, T., Guo, L., ... & Shum, D. (2008).
11
12 Prospective memory in schizophrenia: Further clarification of nature of
13
14 impairment. *Schizophrenia Research* 105(1), 114–124.
15
16
17
18 Wilson, B. A., Cockburn, J., & Baddeley, A. (2003). *The Rivermead Behavioral Memory Test*
19
20 – II. Bury St. Edmunds, UK: Thames Valley Test Co.
21
22 Wilson, B. A., Emslie, H. C., & Foley, J. A. (2004). A new test of prospective memory: The
23
24 CAMPROMPT. *Brain Impairment* 5, 44.
25
26
27 Wilson, B. A., Emslie, H., Foley, J., Shiel, A., & Watson, P. H. K. (2005). *The Cambridge*
28
29 *Prospective Memory Test*. London, U.K.: Harcourt.
30
31 Woods, S. P., Twamley, E. W., Dawson, M. S., Narvaez, J. M., & Jeste, D. V. (2007).
32
33 Deficits in cue detection and intention retrieval underlie prospective memory
34
35 impairment in schizophrenia. *Schizophrenia Research* 90(1), 344–350.
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1. Intra-class correlation (ICC) and 95% confidence interval (CI) for quality of translation and test-retest reliability of VRPMT-CV.

Agreement measures in translation				
	ICC (p value)		95% CI	
Equivalence	0.72, $p < 0.001$		(0.62 to 0.88)	
Clarity	0.63, $p < 0.001$		(0.40 to 0.79)	
Relevance	0.58, $p < 0.001$		(0.35 to 0.77)	
Agreement measures in test-retest reliability				
	ICC	95% CI	ICC	95% CI
	Schizophrenia		Healthy control	
Event-based	0.72, $p = 0.019$	(0.33 to 0.41)	0.903, $p < 0.001$	(0.80 to 0.94)
Time-based	0.79, $p = 0.008$	(0.56 to 0.84)	0.885, $p < 0.001$	(0.75 to 0.92)
Total	0.78, $p = 0.006$	(0.58 to 0.94)	0.915, $p < 0.001$	(0.85 to 0.97)

Table 2. Demographic information and PM performances across schizophrenia and healthy control groups.

		Schizophrenia group		Healthy control subjects		
		(n = 28)		(n = 42)		
		Mean	Standard Deviation	Mean	Standard Deviation	p- value ^a
Age		40.29	(6.91)	36.79	(7.36)	0.54
Years of education		10.57	(2.39)	12.43	(2.32)	0.02
MMSE-CV#		27.64	(3.02)	28.40	(2.04)	0.74
VRPMT-CV## total score		2.29	(1.90)	5.18	(1.02)	< 0.001
Event-based score		1.36	(1.25)	2.75	(0.44)	< 0.001
Time-based score		0.93	(1.12)	2.43	(0.84)	< 0.001
Total CAMPROMPT-CV score		17.96	(9.71)			
Event-based score		9.93	(4.41)			
Time-based score		8.04	(5.88)			
		n	%	n	%	
Gender	Male	23	82.14	20	71.42	0.82
	Female	5	17.86	8	28.58	

^a p values were analysed using Mann-Whitney U test.

MMSE-CV= Chinese version of the Mini-Mental State Examination.

VRPMT-CV =Chinese version of the virtual-reality prospective memory test.

Table 3. Correlation between VRPMT-CV and CAMPROMPT

VRPMT- CV	CAMPROMPT- CV	
	Schizophrenia	Healthy control
Event-based	0.75, $p < 0.001$	0.78, $p < 0.001$
Time-based	0.81, $p < 0.001$	0.79, $p < 0.001$
Total	0.88, $p < 0.001$	0.92, $p < 0.001$

FIGURE 1a. Instructions for completing time-based PM tasks in the VRPMT-CV.

FIGURE 1b. Instructions for completing event-based PM tasks in the VRPMT-CV.

Figure 2. Estimated marginal means of VRPMT-CV total score.

Figure 3. Receiver operating characteristicROC curve of VRPMT-CV.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Acknowledgement

The authors would like to thank staff members of Halfway House III and Joyous Place Hostel of New Life Psychiatric Rehabilitation Association, TWGHs Yeung Sing Memorial Long Stay Care Home and Caritas Jockey Club Lai King Rehabilitation Centre, for providing staunch support in our research project.

For Peer Review Only

Contributors

David Man and David Shum designed the study. David Man and Calvin Yip conducted the statistical analysis. Sarah Tsang, Christina Lee, Janice Young and Pan Yu administered the VRPMT to the participants. David Man wrote the first draft, assisted by David Shum and Calvin Yip. All authors contribute to and have approved the final manuscript.