Validation of the virtual-reality prospective memory test (VRPMT) for individuals with first-episode schizophrenia

David W.K. Man[a], Calvin C.K. Yip[a] ,Christina O.P. Lee[b], Sarah Y. L. Tsang[c], Pan W.P. Yu[d], Janice G.E.Young[e], David H. K. Shum[f]

[a] Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hung Hom, Hong Kong
[b] Hope Community Rehabilitation Day Centre, Haven of Hope Christian Service, Hong Kong
[c] Placidity Place (ICCMW), The Society of Rehabilitation and Crime Prevention, Hong Kong.
[d] The Tsung Tsin Mission of Hong Kong Social Service, Hong Kong
[e] Mercy Medirehab Professional Group Limited
[f] Behavioural Basis of Health Program, Griffith Health Institute, Griffith University, Gold Coast, Australia

* Correspondence should be addressed to: 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
Author title: Professor

Abstract

Objectives: To examine the psychometric properties of a Virtual-Reality Prospective Memory

Test (VRPMT) in individuals experiencing first-episode schizophrenia.

Methods: The VRPMT was administered to 28 individuals with schizophrenia; the test was administered again after one week, to establish test-retest reliability. The concurrent validity of the VRPMT was evaluated by examining correlations between VRPMT score and score on the Chinese version of the Cambridge Prospective Memory Test (CAMPROMPT-CV). The VRPMT performance of individuals with schizophrenia was compared with that of 42 healthy controls to examine the sensitivity and specificity of the test.

Results: The intra-class correlation for test-retest reliability of the total VRPMT–CV score was 0.78 ($p = 0.005$). There was a significant correlation between total VRPMT-CV score and the total CAMPROMPT-CV score ($r = 0.90$, $p < 0.001$). Comparison with healthy controls revealed that total VRPMT-CV score was a sensitive (92.9%) and specific (75%) measure of prospective memory (PM) deficits in individuals with schizophrenia.

Conclusion: VRPMT-CV is a PM assessment with good construct validity, test-retest reliability, sensitivity and specificity in the context of first-episode schizophrenia.

(176 words)

1. Introduction

People with schizophrenia are reported to have deficits in executive functions [1], social cognition [2] and other cognitive processes [3,4]. Their prospective memory (PM) deficits of is one of the cognitive impairments that has attracted increasing attention due to their associated functional implications. These deficits have been found to affect scores on the Activities of Daily Living (ADL) test, the Instrumental Activities of Daily Living (IADL) test and work and social life. PM is the memory required to perform an intended action at some point the future, in response to an external event or at a specific time [5]. PM tasks are classified into event- (e.g. buying a book when passing by a bookshop), time- (e.g. attending an appointment at 11 a.m.) and activity-based (e.g. taking medication after finishing a meal) [6] tasks. PM involves many cognitive processes including retrospective memory (RM), initiation of action, planning, decision-making, inhibitory control and attention [5,7,8]. Individuals with schizophrenia have impaired PM performance [9], especially on time-based tasks [10]. PM problems can lead to poor medication adherence, which in turn may result in relapse and re-hospitalisation [11]. In addition, PM problems have been found to predict community living skills in people with schizophrenia [12]. Early and accurate assessment of PM and early interventions are therefore considered crucial to the rehabilitation of patients with schizophrenia.

PM impairments have been assessed in traumatic brain injury [13], stroke [14],

schizophrenia [9] and bipolar disorders [15] using self-report questionnaires, experimental tasks and psychometric tests. The Prospective and Retrospective Memory Questionnaire (PRMQ) [16,17] and the Comprehensive Assessment of Prospective Memory Questionnaire (CAPM) [18] are examples of self-report PM assessments, but such questionnaires may not be suitable for assessing PM in individuals with schizophrenia because this patient group have limited insight into their condition [20,21]. Discrepancies between self-report and objective measures of PM have been noted [20]; this highlights the need for more objective, validated tools for assessing PM [5]. Objective, paper-and-pencil tests type of psychometric tests of PM are namely the Rivermead Behavioural Memory Test (RBMT) [22], the Cambridge Prospective Memory Test (CAMPROMPT) [23]; the Memory for Intentions Screening Test (MIST) [24] and the Royal Prince Alfred Prospective Memory Test (RPA-ProMem) [5]. The RBMT comprises three within-session event-based PM tasks [19] but no time-based or long-term PM tasks [10, 25]. The CAMPROMPT has good reliability and validity but there was no relationship between CAMPROMPT performance and performance on short-term or daily PM tasks [26]. The MIST requires 30-40 minutes to administer and can be time-consuming and difficult to incorporate into a standard neuropsychological assessment [5,24]. The RPA-ProMem measures both time- and event-based PM using tasks involving a mixture of short- and long-term retention intervals [5] and it is difficult to monitor task implementation and compliance with test procedures.

Ecological validity is considered crucial in PM assessment. Ecological validity can be conceived in terms of two factors, verisimilitude and veridicality [27]. 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). It would be valuable to develop a VR-based, PM assessment which was ecologically valid and could be used in people with schizophrenia.

It has been suggested that VR-based assessment has the potential to play an increasingly important role in cognitive assessment and rehabilitation [28]. VRs can be divided into immersive and non-immersive environments [29]. Several computer-assisted VR assessments have been shown to be reliable, valid, cost effective [30]. They have been found to minimise social desirability bias. The attractive user interface also reduces users' anxiety levels, improves motivation and is helpful for assessing people with poor verbal communication skills [30, 31]. PM is considered amenable to VR-based assessment and one would expect participants assessed in this way to employ a broad range of neurocognitive functions, much as they would in a natural environment [21]. Examples of VR-based PM assessment included the Virtual Library Task (VLT) [33]; the Test Ecologique de Mémoire Prospective (TEMP) [34], the Virtual Bungalow [35] and the Virtual Reality Prospective Memory Test, which is based on a shopping task (VRPMT) [36].

The VRPMT is the latest VR PM assessment task to be developed and the results of research in traumatic brain injury have been encouraging [36]. Thus one of the aims of this study was to evaluate its utility and suitability for use with individuals with schizophrenia. The VRPMT (English version) is based on a shopping centre scenario in which the distraction task consists of carrying out multiple shopping errands according to a list. It was specifically designed to incorporate more demanding PM tasks than previous VR and conventional computerised PM instruments and contains three time- and three event-based PM tasks. A Chinese version of the VRPMT, the VRPMT-CV, was developed for use in the different linguistic and cultural environment of Hong Kong. This study examined the psychometric properties of the VRPMT-CV, namely its concurrent validity, test-retest reliability and sensitivity and specificity to PM deficits. The study also investigated differences between performance on event- and time-based PM tasks, within and between the first-episode schizophrenia and the healthy control groups. The research hypotheses for individuals with first-episode schizophrenia were:

1. Concurrent validity

   There were no statistically significant differences in correlation coefficients (r) between the PM performance of individuals with first-onset schizophrenia as measured by VRPMT-CV and a standardized, validated paper-and-pencil PM, the Chinese version of the CAMPROMPT (CAMPROMPT-CV).

2. Test-retest reliability

There were no statistically significant differences between the mean PM performance in the initial testing and retesting of VRPMT-CV.

3. Sensitivity and specificity to PM deficits in schizophrenia

There were no statistically significant differences in PM performance between individuals with first-onset schizophrenia and the healthy control groups.

4. There were no statistically significant differences between the event-based and time-based PM performance within and between the schizophrenia and healthy control groups.

2. Method

2.1 Participants

Participants in the first-episode schizophrenia group were recruited from three non-governmental organisations (NGOs), 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; Caritas Jockey Club Lai King Rehabilitation Centre, using convenience sampling,

Participants in the control group were normal, healthy adults recruited from the general

population using convenience and snowball sampling. Attempts were made to match the two groups in terms of age, gender and educational level. The common inclusion criteria for both groups were age between 20 and 50 years old and understanding of spoken and written Cantonese, the Chinese dialect of Hong Kong. All participants in the schizophrenia group had been diagnosed with first-episode schizophrenia by a psychiatrist using DSM-IV and were clinically stable. We used an additional selection criterion to exclude individuals with mild cognitive impairment or early dementia: a score of 18 or more on the Chinese version of Mini-Mental State Examination (MMSE-CV) [37].

Exclusion criteria for both groups were significant medical conditions, epilepsy, auditory or visual impairments and a history of alcohol or substance abuse. Individuals with schizophrenia who had received electroconvulsive therapy in the past year were also excluded. Additional exclusion criteria for the healthy control group were a diagnosis of mental illness and having a first-degree relative with a history of schizophrenia or schizoaffective disorder.

Table 1 summarises the demographic characteristics and PM performance of the schizophrenia and healthy control groups. Group comparisons made using the Mann-Whitney U test and chi-square test of independence showed no significant group differences in age or gender, but there was a significant difference in educational level ($MWU = 225$, $z = -3.04$, $p = 0.02$) due to our limited control over the characteristics of volunteer participants in the healthy control group.

Table 1 about here

2.2 Measures

2.2.1 Mini Mental State Examination- Hong Kong Chinese version [37]

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. The sensitivity was 97.5%, specificity was 97.3%; test-retest reliability was 0.78 and inter-rater reliability was 0.99. Cut-off scores were adjusted according to educational level, and were 22, 20 and 18 for individuals with > 2 years, 1-2 years and < 1 year of education respectively. In this study, all patients had a score of 18 or above.

2.2.3 Chinese version of CAMPROMPT (CAMPROMT-CV)

The CAMPROMPT-CV [38] is one of the few Chinese PM tests available and was used as a reference point to establish the concurrent validity of the VRPMT-CV. It consists of three event- and three time-based PM tasks; paper-and-pencil quizzes and puzzles are used as ongoing distraction tasks, as in the original English version [22]. The quizzes are based on popular local actors and actresses, musicians, major Chinese historical events and geographical locations in China to engage the participants. The word puzzles involve word sequencing, filling in missing words, and completing common Chinese idioms. At the same time participants have to perform three time- and three event-based PM tasks [9]. The three

event-based PM tasks of CAMPROMPT-CV are: ‘when you get to a question containing the word ‘Great Wall’ give a map to the testers’; ‘when the testers tell you that there are “five minutes left” give them a message envelope’; ‘when the tester says “the test is over” remind the tester of the locations of five objects placed in the room’. The time-based tasks are: ‘when there are nine minutes left to the end of the ongoing distraction task skip the page you are on and go to the next page’; ‘when there are seven minutes left to the end of the ongoing distraction task’ remind the tester about the keys; ‘five minutes after the ongoing distraction task ends remind the tester to ring the garage’ [9]. Standardised verbal and written instructions were used to administer this test. All time-based or event-based tasks were scored from 0 to 6 on the basis of eight scoring criteria (see Table 2); the range of total scores for 6 tasks was 0 to 36; higher scores indicate better PM performance. The test took 25 minutes to administer. It has been reported to have good reliability and validity [23] and has been used previously to assess PM function in individuals with schizophrenia [12].

Table 2 about here

Virtual-Reality Prospective Memory Test- Chinese version (VRPMT-CV)

The VRPMT-CV was adapted from the English original, which was developed by a team including the first and last authors of this study and has been used to assess PM in individuals

with TBI [36]. The VRPMT-CV is a non-immersive VR programme and was developed using 3DVIA Virtools game development software. The visual elements are delivered through a PC monitor and the auditory elements are transmitted through the speakers of a computer. The VR environment is a shopping centre containing 20 shops. The graphics are based on data captured in a real shopping centre using a digital camera. The VR includes background noise from the shopping centre to increase the verisimilitude. Participants navigate the shopping centre using the mouse and clickable three icons (shopping centre map, shopping list and mobile phone) located in three corners of the screen.

The test comprises three time- and three event-based PM tasks and the distraction task is to navigate to various shops to obtain items from the shopping list. The digital clock on the mobile phone was always set to 8:00pm at the start of the test. The time-based PM tasks required participants send SMSs 4, 8 and 12 minutes after the start of the test, using the digital clock on the mobile phone to keep track of time. Figure 1 shows how a time-based PM task was executed [36]. The event-based PM tasks required participants o press the 'T' key in response to three sales announcements which 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 an announcement attracted one point (maximum event-based PM score = 3). Figure 1b shows how an event-based PM task was executed [36]. Before the test phase participants completed a 14-minute tutorial to familiarise them with the

environment and tasks. A detailed performance summary is generated automatically by the VRPMT programme at the end of a test; the summary includes scores on the time- and event-based PM tasks, number of items bought and frequency with which the participant checked the time.

Figure 1 about here

2.3 Procedures

Approval was given by the Ethics Committee of the Hong Kong Polytechnic University. All participants provided written consent before the assessments were conducted. All participants were offered an incentive of HK$30 per session.

2.3.1 Phase I: Linguistic and cultural validity of the translated version

Before the study proper we involved an expert panel to ensure that the VRPMT-CV was accurately translated and appropriately adapted from the English language original. Written instructions and auditory scripts were translated from English to Cantonese with adaptations to local culture, such as replacing references to the film star Tom Cruise in two of the event-based PM cues with references to Andy Lau, a Hong Kong film star and changing the format of a car number plate from Australian to Hong Kong format. The translated version was evaluated by a panel of four local experts with at least 15 years of rehabilitation experience. They independently rated the equivalence, clarity and relevance of 102 translated items in

terms of on a six-point Likert scale (1 = strongly disagree, 2 = disagree, 3 = slightly disagree, 4 = slightly agree, 5 = agree, 6 = strongly agree).

2.3.2 Phase II: Pilot testing

This aim of this phase was to explore the suitability of the test format to the target population, make any necessary adjustments to the VR programme and optimise administrative and operational procedures for both PM tests. The results of the pilot phase also prompted us to make some adjustments to logistic procedures and certain features of the VR programme before the main study.

Fourteen people with schizophrenia recruited by convenience sampling from a halfway house of the New Life Psychiatric Association completed the CAMPROMPT-CV and VRPMT-CV. The MMSE-CV was used as a screening test in the pilot phase, to ensure that the global cognitive functioning of participants was sufficient to enable them to undertake the VRPMT-CV.

2.3.3 Phase III: Main study

Twenty-eight individuals with schizophrenia were recruited from the three sources mentioned above. After screening for general cognitive impairment with the MMSE-CV they completed the VRPMT-CV and CAMPROMPT-CV; the order in which the tests were

administered was counterbalanced. All participants were given clear instructions before both tests. They were allowed to complete a maximum of two tutorials before the assessment phase of VRPMT-CV to ensure that they had a good understanding of all task requirements. Participants were given a five to ten minute break between the two tests to minimise the effects of mental fatigue. All participants were retested on the VRPMT-CV one week later. All participants with first-episode schizophrenia were also interviewed after the tests to gain information about their views on the usability and acceptability of this new testing method. A group of forty-two healthy participants was recruited and completed the same tests in a similar time frame.

2.4 Data analysis

As a preliminary step all data were screened to determine whether they met the criteria for parametric analysis, appropriate non parametric tests were used where necessary. Two-tailed tests with a significance level of $p < 0.05$ were used in all analyses.

Two-way random effects model of intraclass correlation coefficient (ICC), which took into account of both correlation of scores and agreement of scores, was used to examine the inter-rater reliability of the three areas (equivalence, clarity and relevancy) of all translated items. The items were independently scored by four expert panel members.

The concurrent validity of the VRPMT-CV in the schizophrenia group was assessed by calculating Spearman's rho for correlations between time-based, event-based and total scores on the VRPMT-CV and CAMPROMPT-CV. A two-way mixed effects model of ICC was used to determine the test-retest reliability of the VRPMT-CV. Time-based, event-based and total scores on the VRPMT-CV were obtained at two time points (first test and retest) from twenty-eight 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) ANOVA with years of education as a covariate; other demographic variables such as age was similar in the two groups and were therefore not included in the analysis. The sensitivity and specificity of VRPMT-CV to PM deficits in schizophrenia was evaluated using receiver operating characteristic (ROC) curve analysis. The area under the curve (AUC) was calculated to determine the usefulness of the diagnostic value of the ROC curve. Mann-Whitney U tests were used to assess group differences in number of time checks and number of shopping tasks completed. Moreover, Mann-Whitney U test was used to examine differences in computer-generated data in number of time checking and number of shopping tasks completed between the two participant groups.

3. Results

3.1 Quality of translation

Translated items were divided into two main categories: those related to the 6 PM tasks (time-based and event based) and those related to the distraction task and environment (e.g. announcements, signs, names of shops and pop-up windows showing messages and response options). Agreement between the four expert panel members was moderate with respect to all aspects of the translation: equivalence ICC(2,2) = 0.72, $p < 0.001$; clarity ICC(2,2) = 0.63, $p < 0.001$ and relevance ICC(2,2) = 0.58, $p < 0.001$. However average scores for all three aspects of all translated items were at least 4 (i.e. slightly agree, agree or strongly agree), indicating a high 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 is culturally relevant).

3.2 Group Comparison on PM performance

PM was divided into event- and time based components. An ANOVA with group as between-subject factor and task type as a within-subjects factors indicated that performance on the two types of PM task was similar[F(1,52) = 0.08, p = 0.093], and there was no group by task type interaction, [F(1,52) = 0.913, p = 0.344]. But in between-subjects testing, there were significant differences between the schizophrenia and health control groups in their event- and time-based PM performance [F(1,52) = 33.52, $p < 0.001$]. In the analysis, both gender [F(1,52) = 0.114, p = 0.737] and level of education [F(1,52) = 0.929, p = 0.340] did not significantly contribute to the differences in their performance in event- and time based PM.

In further sub-group analysis, it was found that for both groups, there were no statistically significant differences between time-based and event-based PM performance, with p=0.121 for schizophrenia group and p=0.059 for health group respectively. The number of shopping completed in the schizophrenia group (mean = 4.46) was shown to be significantly lower than the healthy group by Mann-Whitney U test (mean = 9.93; $p < 0.001$). The number of time-checking in schizophrenia group (mean = 4.89) was also significantly less than the healthy control group (mean = 9.86; < 0.001).

Figure 2 about here

3.3 Concurrent validity

In the healthy group, the VRST-CV scores correlated with CAMPROMPT-CV scores: for event-based score, $r = 0.782$, $p<0.001$; for time-based score, $r = 0.797$, $p<0.001$ and for total score, $r = 0.924$, $p<0001$.

In the schizophrenia group VRST-CV scores were positively correlated with CAMPROMPT-CV scores, total event-based score: $r = 0.758$, $p < 0.001$; total time-based score: $r = 0.815$, $p < 0.001$; total score: $r=0.889$, $p< 0.001$.

3.4 Test-retest reliability

In the healthy group, the test-retest reliability of the VRPMT-CV scores showed good

reliability in time-based, event-based and total scores between the two time points. For time-based score, ICC (3,1) was found to be 0.885, $p < 0.001$; for event-based score, ICC (3,1) was 0.903, $p<0.001$, and for total score, ICC (3,1) was 0.915, $p<0.001$.

In the schizophrenia group VRPMT-CV initial test and re-test performances were positively correlated with respect to the time-based tasks, ICC (3,1) = 0.79, $p = 0.008$, and total performance (sum of scores on both time- and event-based tasks), ICC (3,1) = 0.78, $p = 0.0058$, but not event-based tasks, ICC (3,1) = 0.32, $p = 0.19$. On the VRPMT-CV re-test patients performed better on the event-based PM tasks, with scores being on average 0.5 points higher than in the first test. These results indicated that the VRPMT-CV has good test-retest reliability with respect to time-based PM tasks but not event-based PM tasks in this sample of patients.

3.5 Sensitivity and specificity

In the schizophrenia group, 64.3% got CAMPROMPT-CV scores which were below the cut off, indicating that they were suffered from PM problems. However, in the healthy control group, only 7.1% suffered from PM problems. Using the cut off scores of CAMPROMPT-CV as a golden standard for PM problems, the receiver operating characteristic (ROC) curve of the VRPMT-CV was thus analysed.

Using the cut off 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 area under the curve

(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 in VRPMT-CV time-based PM, the sensitivity and specificity were 83.3% and 80% respectively, and the AUC was 0.854, $p<0.001$. In order to increase the specificity of the time-based and event-based PM, the cut off points of the VRPMT-CV event-based score and time-based score were increased. When we adopted a cut off at 2.5 for event-based PM, the sensitivity and specificity were change to 75% and 90% respectively. For time-based score, using an increased cut off of 2.5, the sensitivity and specificity were changed to 52.8% and 95% respectively. Therefore, the time-based and event-based scores of VRPMT-CV could have a higher specificity as a trade-off of decreasing its sensitivity (Figure 3).

Among the three AUC of ROC curves, the AUC of the total score of VRPMT-CV had the largest area. This result indicated 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 a better index for screening PM impairment in individuals with schizophrenia.

4. Discussion

Individuals with schizophrenia performed worse on the VRPM-CV than the healthy control group. This finding is consistent with previous research that suggests PM impairment is a significant cognitive problem in schizophrenia and affects the everyday functioning of

people with schizophrenia [39]. Wang and associates [39] and Wood and associates [40] found that PM impairments mainly originated in problems with cue detection and intention retrieval. These problems may be due to the working memory impairments in individuals with schizophrenia. It has also been argued that cue detection and intention retrieval depend on the ability to self-initiate actions; prefrontal lobes are involved in self-initiation and prefrontal abnormalities have been reported in patients with schizophrenia [39]. Thus strategies to enhance visual attention and working memory, provision of emotional cues, formation of implementation intentions or use of goal-directed future thinking could be used to improve the PM of people with schizophrenia.

We did not observe the expected difference in performance on time- and event-based PM tasks. This might be because there were too few test items or because the scoring system is not sufficiently sensitive. The small range of the subscore (0 to 3) is a possible reason. It should be noted that there was a trend towards better performance on the event-based tasks, suggesting that the time based tasks more difficult, perhaps because they are more abstract, less dependent on retrospective memory and successful performance cannot be based on external cues [8,10,39]. Moreover it has been shown that individuals with schizophrenia have time processing deficits which may have a greater impact on time-based PM than event-based PM [39]. People with schizophrenia should thus be encouraged to use feedback or reminders on tasks, time monitoring and self-initiation strategies to cope with the PM problem [42].

The hypothesis that in patients with schizophrenia, performance on the VRPMT-CV would correlate with performance on the CAMPROMPT-CV, a standardised, validated paper-and-pencil measure of PM was supported, providing evidence for the concurrent validity of the VRPMT-CV.

Sensitivity analysis suggested that schizophrenic patients had worse PM than healthy controls when PM was assessed in terms of number of correct responses on event- and time-based tasks in the VRPMT-CV [39].

The hypothesis that patients with schizophrenia would perform similarly on the VRPMT-CV in the first test and re-test was partially supported. Total score on the time-based tasks showed good test-retest reliability, but test-retest reliability for total score on the event-based tasks was poor; in general patients performed better on the event-based tasks in the retest. This may reflect a practice effect limited to event-based tasks as these required the participant to respond to sales announcements and withhold responding to distracting announcements. As the number and content of these announcements were unchanged on the retest, participants might easily have learned something about them which facilitated performance on the event-based tasks thus lowering the test-retest reliability of this component of the VRPMT-CV.

The VRPMT-CV has several positive features. It is a PM assessment tool that can be used on any standard personal computer, notebook or tablet PC and is thus handy, mobile and

convenient. The result of an assessment is automatically generated by the programme which runs the test, so scoring requires minimal time and effort and is accurate, reliable and objective. The VRPMT-CV is more ecologically valid than traditional paper-and-pencil PM assessments because the virtual environment is a familiar one and the PM tasks are part of most participants' everyday experience. Another study has shown that results of VR assessments correlate significantly with everyday functioning [36]. An important advantage of the VRPMT-CV is that it is easier for people with poor communication or social skills, such as individuals with first-episode schizophrenia, to carry out, because stress and anxiety associated with communicating with the assessor is greatly reduced.

Qualitative feedback on the tests from participants suggested that the VRPMT-CV had good usability. Participants found the test interesting and attractive as it was similar to a computer game. The activity increased their attention to the tasks and their motivation to perform well. Some participants reported that the test was user-friendly. Participants also found the graphics and sales announcements familiar and appealing, as they were culturally relevant and simulated the environment of a real-life shopping centre.

The current study has some limitations. Firstly, the psychometric properties of the new translation of the instrument have not been fully studied in the healthy population and it should be prioritized before looking at clinical presentations. Secondly, we did not record detailed information about patient medication such as chlorpromazine-equivalent dosages of

antipsychotic medication or the types of medication patients were receiving at the time of this study. This may give rise to questions about whether performance was influenced by side-effects of medication. Thirdly, the NGOs had difficulty recruiting individuals with schizophrenia who fulfilled the inclusion and exclusion criteria for the study so the sample was smaller than we hoped. Fourthly, even after completing two tutorial sessions some of the participants in the schizophrenia group required additional explanations about how to do the VRPMT-CV and variations in the assessors' additional explanations may have influenced performance in these cases. Lastly, there is not alternate version of the VRPMT-CV to minimise practice effects in cases where reassessment is required. This was a more obvious shortcoming with respect to assessment of event-based PM.

We recommend a number of modifications to the VRPMT-CV. Firstly, the tutorial should include auditory as well as written instructions to make it easier for participants to understand the test procedure and task requirements and reduce the role of assessors' supplementary instructions which are a potential source of variance in performance. Secondly, the content or number of auditory announcements in the VRPMT-CV should be modified or increased to minimise practice effects and increase the test-retest reliability of the event-based component. Lastly, development of alternative versions based on other real-life scenarios or using a different distraction tasks might improve the ability of the test to predict PM performance in a variety of real-life situations as well as reducing practice effects.

5. Conclusion

This study showed that the VRPMT-CV has good concurrent validity and is a valid, objective and convenient measure of PM performance in people with first-episode schizophrenia. The VRPMT-CV has good test-retest reliability, with the exception of the event-based PM component. It was found to be a sensitive method of assessing PM deficits in clinical settings. Changes to the content and structure of the VRPMT-CV can be made in future developmental work.

Conflict of interest

None of the authors have any conflict of interest affecting the conduct of this study or the preparation of this manuscript.

Acknowledgment

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 study design, in the collection and interpretation of data, in writing the report or in the decision to submit the paper for publication.

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.

References

[1] Walker E, Kestler L, Bollini A, Hochman KM. Schizophrenia: etiology and course. Ann Rev Psychol 2004; 55:401-30.

[2] Goldberg JO, Schmidt LA. Shyness, sociability, and social dysfunction in schizophrenia. Schizophr Res 2001; 48: 343-9.

[3] Bowie CR, Harvey PD. Treatment of cognitive deficits in schizophrenia. Curr Opin Investig Drugs 2006;7: 608-13.

[4] Heinrichs RW, Zakzanis KK. Neurocognitive deficit in schizophrenia: a quantitative review of the evidence. Neuropsychology 1998;12: 426-45.

[5] Radford KA, Lah S, Say MJ, Miller LA. Validation of a new measure of prospective memory: The Royal Prince Alfred Prospective Memory Test. Clin Neuropsychol 2011;25 (1): 127-40.

[6] Ellis J. Prospective memory or the realization of delayed intentions: a conceptual framework for research. In: Brandimonte M, Einstein GO, McDaniel MA, editors. Prospective memory: theory and applications. Mahwah, NJ: Lawrence Erlbaum Associates Inc; 1996, p. 1-22.

[7] Fish J, Wilson BA, Manly T. The assessment and rehabilitation of prospective memory problems in people with neurological disorders: a review. Neuropsychol Rehabil 2010; 20 (2):161-79.

[8] Groot Y, Wilson B, Evans J, Watson P. Prospective memory functioning in people with and without brain injury. J Int Neuropsychol Soc 2002;8:645-54.

[9] Au RCW, Ungvari GS, Lee E, Man D, Shum D, Xiang YT, Tang WS. Prospective memory impairment and its implications for community living skills in bipolar disorders. Bipolar Disord 2013; 15 (8):885-92.

[10] Shum D, Ungvari GS, Tang WK, Leung JP. Performance of schizophrenia patients on time-, event-, and activity-based prospective memory tasks. Schizophr Bull 2004;30: 693-701.

[11] Moore A, Sellwood W, Stirling J. Compliance and psychological resistance in schizophrenia. Br J Clin Psychol 2000;39:287-95.

[12] Au RWC, Man D, Shum D, Lee E, Xiang YT, Ungvari GS, Tang W.K. Assessment of prospective memory in schizophrenia using the Chinese version of the Cambridge Prospective Memory Test: a controlled study. Asia Pac Psychiatry 2014; 6:54-61.

[13] Shum D, Levin H, Chan RC. Prospective memory in patients with closed head injury: a review. Neuropsychologia 2011;49:2156–65.

[14] Brooks B M, Rose FD, Potter J, Jayawardena S, Morling A. Assessing stroke patients' prospective memory using virtual reality. Brain Inj 2004:18 (4):391-40.

[15] Lee E, Xiang YT, Man D, Au RWC, Shum D, Tang WK, Chiu HFK, Wong P, Ungvari GS. Prospective memory deficits in Chinese patients with bipolar disorder. Arch Clin Neuropsychol 2010;25:640-7.

[16] Hsu YH, Hua MS. Taiwan version of the Prospective and Retrospective Memory Questionnaire: latent structure and normative data. Arch Clin Neuropsychol 2011; 26 (3):240-9.

[17] Smith G, Della Sala S, Logie R, Mayor EA. Prospective and retrospective memory in normal aging and dementia: a questionnaire study. Memory 2000;8:311-21

[18] Chan RCK, Qing YH, Wu QP, Shum, D. Prospective memory in healthy Chinese people: the latent structure of the Comprehensive Assessment of Prospective Memory Questionnaire. Neuropsychol Rehabil 2010; 20 (3): 459-70.

[19] Roche N L, Fleming JM, Shum DHK. Self-awareness of prospective memory failure in adults with traumatic brain injury. Brain Inj 2002; 16:931-45.

[20] Chan RCK, Wang Y, Ma Z, Hong XH, YuanY, Yu X, Li Z, Shum D, Gong QY. Objective measures of prospective memory do not correlate with subjective complaints in schizophrenia. Schizophr Res 2008;103 (1-3):229-39.

[21] Kurtz MM, Baker E, Pearlson GD, Astur RS. A virtual reality apartment as a measure of medication management skills in patients with schizophrenia: a pilot study. Schizophr Bull 2007;33 (5):1162–70.

[22] Wilson BA, Emslie HC, Foley JA. A new test of prospective memory: the CAMPROMPT. J Int Neuropsychol Soc 2004;10:44.

[23] Wilson BA, Emslie, HC, Foley J, Shiel A, Watson P, Hawkins K et al. The Cambridge Prospective Memory Test. London: Harcourt; 2005.

[24] Raskin S. Memory for intentions screening test. J Int Neuropsychol Soc 2004:10

(Suppl. 1): 110.

[25] Mathias J, Mansfield K. Prospective and declarative memory problems following moderate and severe traumatic brain injury. Brain Inj 2005;19: 271-82.

[26] Fish J, Evans J, Nimmo M, Martin E, Kersel D, Bateman A. et al. Rehabilitation of executive dysfunction following brain injury: ‘‘content-free’’ cueing improves everyday prospective memory performance. Neuropsychologia 2007; 45:1318-30

[27] Chaytor N, Schmitter-Edgecombe M. The ecological validity of neuropsychological tests. Neuropsychol Rev 2003; 13: 18-97.

[28] Rose FD, Brooks BM, Rizzo A. Virtual reality in brain damage rehabilitation: review. Cyberpsychol Behav 2005;8 (3):241-62.

[29] Man DWK, Chung JCC, Lee GYY. Evaluation of a virtual reality-based memory training programme for Hong Kong Chinese older adults with questionable dementia: a pilot study. Int J Geriatr Psychiatry 2012; 27: 513-20.

[30] Veling, W., Moritz, S., van der Gragg M. Brave new world- review and update on virtual assessment and treatment in psychosis. Schizophr Bull 201;40(6):1194-97.

[31] Kim K, Kim SI, Cha KR, Park J, Rosethal MZ, Kim JJ, Han K, Kim IY, Kim CH. Development of a computer-based behavioral assessment of checking behavior in obsessive-compulsive disorder. Compr Psychiat 2010;51:86-93.

[32] Man DWK, Chung JCC, Mak MKY. Development and validation of the Online Rivermead Behavioral Memory Test (OL-RBMT) for people with stroke. NeuroRehabilitation 2009; 24: 231-6.

[33] Renison B, Ponsford J, Testa R, Richardson B, Brownfield K. The ecological and construct validity of a newly developed measure of executive function: the Virtual Library Task. J Int Neuropsychol Soc 2012;18: 440–450.

[34] Potvin M J, Rouleau I, Audy J Charbonneau S, Giguere JF. Ecological prospective memory assessment in patients with traumatic brain injury. Brain Inj 2010; 25:192-205.

[35] Attree EA, Dancey C P, Pope A L. An assessment of prospective memory retrieval in women with chronic fatigue syndrome using a virtual-reality environment: an initial study. Cyberpsychol Behav 2009; 12 (4): 379-85.

[36] Canty AL, Felming J, Patterson F, Green H, Man D, Shum D. Evaluation of a virtual reality prospective memory task for use with individuals with severe traumatic brain injury. Neuropsychol Rehabil 2014;24 (2):238-65.

[37] 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 J. Psychiatry 4 (2), 25-28.

[38] Lou ZL, Dou Z L, Zheng JL, Man DWK. The study of the Chinese version of Cambridge Prospective Memory Test (CAMPROMPT) for traumatic brain injury (unpublished Master thesis). Guangzhou: Sun Yat Sen University; 2009.

[39] Wang Y, Chan RC K, Hong X, Ma Z, Yang T, Guo L et al. Prospective memory in schizophrenia: further clarification of nature of impairment. Schizophr Res 2008;105: 114–24.

[40] Woods SP, Twamley EW, Dawson MA, Narvaez JM, Jeste DV. Deficits in cue detection and intention retrieval underlie prospective memory impairment in schizophrenia. Schizophr Res 2007;90 (1): 344-50.

[41] Peterburs J, Nitsch AM, Miltner WHR, Straube T. Impaired representation of time in schizophrenia is linked to positive symptoms and cognitive demand. PLoS ONE 2013; 8(6): e67615. doi:10.1371/journal.pone.0067615

[42] D'Ydewalle G, Luwel K, Brunfaut E. The importance of ongoing concurrent activities as a function of age in time- and event-based prospective memory. Eur J Cog Psychol 1999; 11: 219-37.

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 (n = 28)		Healthy control subjects (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

	CAMPROMPT- CV	
VRPMT- 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, $p < 0001$

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

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

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

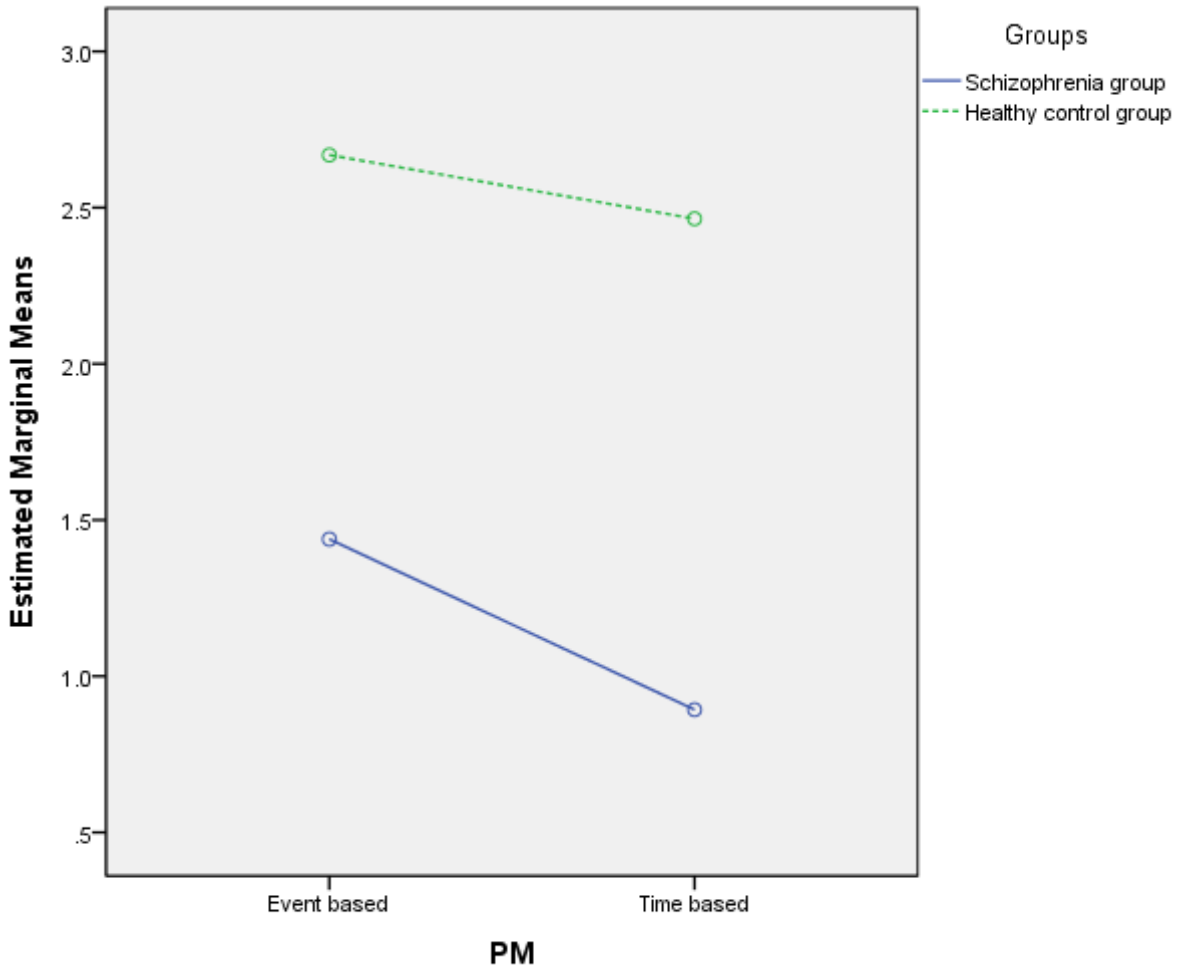

Figure 3. Receiver-operating characteristic curve of VRPMT-CV.