Published in final edited form as:

Schizophr Res. 2016 August ; 175(1-3): 90–96. doi:10.1016/j.schres.2016.03.038.

Validation of a Computerized Test of Functional Capacity

Richard S.E. Keefe, Ph.D.[a,b], **Vicki G. Davis, Dr.PH.**[b], **Alexandra S. Atkins, Ph.D.**[b], **Adam Vaughan, Ph.D.**[b], **Tom Patterson, Ph.D.**[c], **Meera Narasimhan, M.D.**[d], and **Philip D. Harvey, Ph.D.**[e,f]

[a]Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC, USA

[b]NeuroCog Trials, Inc., Durham, NC, USA

[c]Department of Psychiatry, University of California San Diego, San Diego, CA, USA

[d]Department of Neuropsychiatry, University of South Carolina, Columbia, SC

[e]Department of Psychiatry and Behavioral Sciences, University of Miami School of Medicine, Miami, FL, USA

[f]Research Service, Bruce W. Carter VA Medical Center, Miami, FL

Abstract

Regulatory guidance for schizophrenia cognition clinical trials requires that the assessment of cognitive change is accompanied by a functionally meaningful endpoint. However, currently available measures are challenged by resistance to change, psychometric weaknesses, and for interview-based assessments, dependence upon the presence of an informant. The aims of the current study were to: 1) assess the validity, sensitivity, and reliability of the Virtual Reality Functional Capacity Assessment Tool (VRFCAT) as a measure of functional capacity; 2) determine the association between performance on the VRFCAT and performance on the MATRICS Consensus Cognitive Battery (MCCB); and 3) compare the metrics of the VRFCAT with the UCSD Performance-based Skills Assessment (UPSA). 167 patients with schizophrenia and 166 healthy controls completed the VRFCAT, UPSA, and the MCCB at baseline. The VRFCAT and UPSA were completed again at follow-up. The VRFCAT, MCCB, and UPSA were very sensitive to impairment in schizophrenia (d = 1.16 to 1.22). High test-retest reliability was demonstrated for VRFCAT total completion time and the UPSA total score in patients (ICC=0.81 and 0.78, respectively). The UPSA demonstrated significant practice effects in patients (d=0.35),

Corresponding author: Richard S.E. Keefe, Ph.D. Professor of Psychiatry, Duke University Medical Center, Box 3270, Durham, NC 27710, Phone: (919) 684-4306, Fax: (919) 684-2632, richard.keefe@duke.edu.

Contributors

RSE Keefe directed the study, obtained the funding, wrote and edited the protocol, interpreted the analyses, and wrote and edited the paper. VM Davis performed and interpreted the statistical analyses and edited the paper. AM Atkins obtained the funding, wrote and edited the protocol, interpreted the analyses, and wrote and edited the paper. A Vaughan managed the study and edited the paper. M Narasimhan supervised the collection of data, and edited the paper. T Patterson supervised the collection of data and edited the paper. PD Harvey supervised the collection of data, edited the protocol, and wrote and edited the paper.

while the VRFCAT did not (d=−0.04). VRFCAT total completion time was correlated with both UPSA (r=−0.56, p<0.0001 for patients and −0.58, p<0.0001 for controls) and MCCB Composite (r=−0.57, p<0.0001 for patients and −0.68, p<0.0001 for controls). The VRFCAT is a highly reliable and sensitive measure of functional capacity with associations to the UPSA and MCCB. These results provide encouraging support for a computerized functional capacity assessment for use in schizophrenia.

Keywords

Functional capacity; Cognition; Virtual reality; Computerized testing; Everyday functioning

1. Introduction

Schizophrenia is a serious mental illness accompanied by cognitive impairment that persists over the lifetime (Bilder et al., 2000; Harvey et al., 1999; Saykin et al., 1994; Seidman et al., 2010). Cognitive impairment is a substantial determinant of the adverse functional consequences of the disorder, such as the inability to live independently, maintain interpersonal relationships, and work (Bowie & Harvey, 2005; Leifker, 2009). Unfortunately, current antipsychotics provide only minimal cognitive benefit (Davidson et al., 2009; Keefe, 2007). Substantial research is focused on attempting to discover treatments aimed at improvement of cognitive functioning for patients with schizophrenia. A consensus initiative, the Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) project (Marder & Fenton, 2004), was aimed at the unmet need for standardized methods for development of treatments of cognitive impairment associated with schizophrenia. During this work, the FDA representatives stated their position, consistent with earlier guidance (Laughren, 2001) that cognitive improvements alone are not a sufficient demonstration of treatment efficacy and that approval of new treatments for cognitive impairments would also require evidence that any cognitive improvements were clinically meaningful. Thus, in the US and likely elsewhere, clinical trials for cognitive enhancement in schizophrenia must demonstrate improvement on a standard cognitive measure and improvement on a measure aimed at indexing meaningful current or potential benefit to the patient's functioning. This is known as a 'co-primary requirement'.

The MATRICS project leaders made very specific recommendations about which standard cognitive measures would validly measure cognitive change, and developed a battery of tests for this purpose known as the MATRICS Consensus Cognitive Battery (Nuechterlein et al., 2008). However, they recommended that a change in real-world functioning should not be a requirement of drug approval. Their reasoning was that there are many elements of real-world functioning in schizophrenia, such as maintaining employment, living independently, and maintaining social relationships. However, changes in these aspects of real-world functioning might not be observed in treatment studies because they are likely to take longer to occur, even in the presence of cognitive gains, than the duration of a typical clinical trial. Furthermore, real-world functional change is dependent upon a variety of circumstances unrelated to treatment, such as the local and national economies, state funding, and whether a patient is receiving disability payments (Rosenheck et al., 2006). The MATRICS group

thus recommended that clinical meaningfulness could be demonstrated through assessment of the potential to demonstrate real-world functional improvements associated with cognitive change (Green et al, 2008). Such outcomes are currently referred to as measures of "functional capacity" (Moore et al., 2007). These measures are important not only for treatment development, but also for the collection of outcome data and clinical treatment response.

The Validation of Intermediate Measures study (Green et al., 2011) conducted a head to head comparison of several interview and performance-based assessments of functional capacity. Interview-based assessments solely reliant on patient self-report were found to manifest minimal correlation with performance on the MATRICS Consensus Cognitive battery (MCCB). Other studies using high contact informants have found higher cross-sectional correlations and greater sensitivity to treatment (Durand et al., 2015; Keefe et al., 2015a; Sabbag et al, 2011). However, many patients do not have someone who could reliably observe cognitively relevant behaviors, which could prevent enrollment of such patients in clinical trials. In contrast, performance-based assessments of functional capacity do not require informants. Further, the test most frequently used to measure functional capacity in schizophrenia clinical trials, the UCSD performance based skills assessment (UPSA), has been found repeatedly to manifest substantial cross-sectional correlations with cognition (Leifker et al., 2011).

The UPSA and other previous performance-based functional capacity measures have typically involved pen and paper methods, with requirements for patients to role-play the performance of certain tasks. Some of the critical tasks in these assessments also require functional activities that are becoming outdated (calling directory assistance or writing a paper check) or that are easily learned and susceptible to practice effects because there is only one form of the test. Some components of the tasks are challenging to validly adapt to cross national/cross-cultural trial designs (Velligan et al, 2012). As the UPSA is designed to detect disability and not to generate a normal distribution of scores referenced to healthy performance, there is also the potential for a proportion of patients to obtain valid baseline scores that are so high that they cannot improve with treatment.

In an attempt to enhance the assessment of functional capacity, we have developed a computerized, immersive (i.e., generating a three-dimensional image that appears to surround the user), and potentially remotely deliverable assessment referred to as the Virtual Reality Functional Capacity Assessment Tool (Atkins et al, 2015; Ruse et al., 2014). The VRFCAT consists of 6 versions of 4 mini scenarios that include navigating a kitchen and planning a trip to the grocery store, catching a bus to and from the store (selecting the correct bus and paying the correct fare), shopping for and purchasing the correct items at the store, and returning home. Thus, this assessment strategy captures several of the domains of other functional capacity measures: transportation, finances, household management, and planning. The alternate forms are a unique feature of this assessment and the scenarios have the potential to be updated and cross-culturally adapted in geographical regions where computer use, public transportation, and grocery stores are common.

The current study evaluated the psychometric characteristics of the VRFCAT, with special reference to correlations with the MCCB and the leading co-primary measures used in schizophrenia clinical trials, the UPSA and the Schizophrenia Cognition Rating Scale (SCoRS). Following the processes by which the MATRICS-CT group evaluated co-primary measures, we evaluated the characteristics of the VRFCAT in terms of the following areas: sensitivity to impairment; test-retest reliability; practice effects; relationship to cognitive impairment; relationship to real-world functioning; and tolerability for patients. We also investigated the relationship of the VRFCAT to key demographics and symptoms, and compared the alternate forms of the VRFCAT for similarity.

2. Experimental/Materials and Methods

2.1 Participants

Schizophrenia patients and healthy controls were recruited at three research sites: 1) The University of South Carolina under the supervision of Dr. Meera Narasimhan; 2) The University of Miami Miller School of Medicine under the supervision of Dr. Philip Harvey; and 3) The University of California, San Diego School of Medicine under the supervision of Dr. Thomas Patterson.

Patients met criteria for DSM-IV TR schizophrenia, any subtype. All patients completed a structured diagnostic interview, administered by a trained interviewer: the Mini International Neuropsychiatric Interview, 6th Edition (Sheehan et al., 1998). Patients with PANSS symptom severity scores greater than 5 (“moderately severe”) on either item P1 or P3 (delusions or hallucinatory behavior) were excluded from the study, in line with the standards for cognitive enhancement clinical trials from the MATRICS initiative (Buchanan et al., 2005; Buchanan et al., 2011). Patients were also screened for their ability to engage in testing. Those who were uncooperative, suffered from extreme cognitive impairment, had another DSM-IV diagnosis that would exclude the diagnosis of schizophrenia, or had severely limited eyesight were excluded. Participants who participated in studies of cognition with any of the same measures within the last 12 months were not included. None of the participants in our previous feasibility study of the (Ruse et al., 2014) were included. Healthy controls were excluded for a lifetime history of a psychotic disorder, bipolar disorder, or major depression. Other exclusionary criteria for both samples included inability to provide personal informed consent, history of brain trauma, documented neurologic disorder, medical conditions interfering with daily functioning, and current or recent substance abuse. Approval from the Institutional Review Board (IRB) was secured at each site and at the Sponsor site and all patients and healthy controls signed an IRB approved consent form.

2.2 VRFCAT Description

The VRFCAT measures four different functional abilities: checking for the availability of items to complete a recipe, taking a bus, shopping in a store, and managing currency. These scenarios use immersive computerized technology. All participants received a brief tutorial, which included sample items similar to those from the test and practice in using the mouse and computer. There were 12 different objectives, presented in Table 1. For each objective

the dependent variables were accuracy of performance and time to completion. For all objectives, participants who were unable to complete the objective within a pre-specified time period were given a time to completion score of 300 seconds for that objective and automatically progressed to the next objective. Progressions were scored in a binary fashion with 1 indicating the subject was progressed to the next objective at least once during the testing period. All subjects completed a VRFCAT questionnaire that asked them to rate pleasantness, ease of use, clarity of instructions and realism of the VRFCAT virtual environment on a 7-point Likert scale. Six different forms of the VRFCAT were developed and tested in this study.

2.3 Convergent Validity Assessments

2.3.1 MATRICS Consensus Cognitive Battery (MCCB)—The MCCB measures seven separable cognitive domains: speed of processing; attention/vigilance; working memory (verbal and nonverbal); verbal learning; visual learning; reasoning and problem solving; and social cognition. Administration of the MCCB requires about 75–90 minutes. The subtests were administered in the standard order. The MCCB scoring program yields seven domain scores and a composite score, which are standardized to the same T-score measurement scale with a mean of 50 and an SD of 10 (Kern et al., 2008).

2.3.2 Functional Capacity—We used the same version of the UPSA that was used in the MATRICS-CT Validation of Intermediate Measures study (UPSA-VIM) (Sabbag et al., 2011). The UPSA-VIM was designed to assess the ability to perform everyday tasks needed for independent community functioning. The UPSA-VIM evaluates five areas: household chores, communication, finance, transportation, and organization/planning. Raw scores from each subtest are transformed to yield comparable scores (ranging from 0 to 20) for each and a summary score ranging from 0 to 100. Higher scores reflect better performance.

2.3.3 Schizophrenia Cognition Rating Scale (SCoRS)—The SCoRS is an interview-based measure of cognitive impairment with questions aimed at the degree to which this impairment affects day-to-day functioning. Each item is rated on a scale ranging from 1–4 with higher scores reflecting a greater degree of impairment, with anchor points for all levels of the 4-point scale. The anchor points for each item focus on the degree of impairment in that ability and the degree to which the deficit impairs day-to-day functioning. Raters considered cognitive deficits only and attempted to rule out non-cognitive sources of the deficits. Final scores for the SCoRS best judgment ratings generated by an interviewer who had administered the scale to the patient and informant (family member, friend, social worker, etc.). A simple sum of the 20 SCoRS items was used as the dependent variable (Keefe et al., 2015a).

2.3.4 Real-World Functional Outcomes—The rating scale employed was the Specific Levels of Functioning (SLOF) (Schneider & Struening, 1983). The original SLOF is a 43-item informant-rated 5-point Likert scale of a person's behavior and functioning that was abbreviated to assess the following domains: Interpersonal Functioning, Everyday Activities, and Vocational Functioning. The dependent variable was the average score for the items in each subscale (range1–5), leading to a summed total score that could range from 3–

15 for each participant, with higher scores indicating better functioning. The same informant who provided information for the SCoRS served as the SLOF informant.

2.3.5 Clinical Symptom Severity Ratings—Severity of psychotic and negative symptoms in patients was assessed using the Positive and Negative Syndrome Scale (Kay et al, 1987). This 30-item scale is widely used to rate clinical symptoms in treatment trials in schizophrenia.

All subjects completed the VRFCAT, UPSA-VIM, and the MCCB at Visit 1. Patients also received SCoRS and SLOF ratings and were rated with the PANSS at visit 1. All subjects who participated in the initial assessment were asked to return within 7 to 14 days of initial testing for a second assessment with the UPSA-VIM and an alternate version of the VRFCAT. Time 1 and time 2 versions of the VRFCAT were randomly assigned to patients and controls in equal numbers across the 6 forms, with no one tested with the same form twice. Key outcome measures for the VRFCAT were the total time to complete all objectives, progressions, and errors. Equal numbers of healthy controls and schizophrenia patients were scheduled to be tested with each form at each assessment.

2.4 Statistical Analyses

The VRFCAT total time, total errors, and progression scores were transformed into standardized T-scores for purposes of analyses. Regression models containing age and gender as predictors were fit to data from the healthy control sample which were trimmed to remove the 4 highest and 4 lowest scores. Predicted scores for each schizophrenia subject were calculated using parameter estimates from these regression models and subtracted from their actual scores to yield a residual value. The sign of the residuals was adjusted so that higher values reflected better performance and were then transformed into T-scores using the SD of the residuals from the regression models on the healthy control sample. The resulting T-scores reflect how each of the schizophrenia subjects performed relative to expectations for a healthy control subject of the same age and gender in this sample. Raw scores for the MCCB subtests were converted to T-scores and a composite score through the use of the MCCB scoring program. Scores on the UPSA-VIM were standardized with a range of 0 to 100. Totals for the SCoRS, SLOF (as modified), and PANSS were used in analyses.

Group comparisons were made with t-tests and the magnitude of the effect size differences was calculated with Cohen’s d. Retest stability was measured with intra-class correlation coefficients (ICC). Test-retest practice effects were determined with Cohen’s d and t-tests. The relationships among the measures were calculated with Pearson correlation coefficients and multiple regression analyses. Since the correlations between individual SCoRS items and individual SLOF domains were based on single-item SLOF Likert-scales, they were based upon Spearman correlation calculations. All other correlations used Pearson correlations calculations.

3. Results

The study sample included 167 patients with schizophrenia (of whom 158 have complete data on all VRFCAT measures) and 166 healthy controls. One extreme outlier (7 SD below

the mean total time) in the healthy control group was eliminated from all analyses. Descriptive characteristics for the resulting 323 participants in the two samples are presented in Table 2. As can be seen in the table, age, sex, primary language, race and ethnic characteristics, and maternal education did not differ across the patient and control samples. Patients had less education, were slightly more likely to rate themselves as less comfortable with the computer, and were more likely to be unemployed.

3.1. Sensitivity to group differences

Table 3 presents performance on the MCCB, the UPSA-VIM, and the T-scores for the VRFCAT variables (total time, total errors, and progression) for both participant samples, and clinician ratings on the SCoRS, SLOF and PANSS. As can be seen in the table, healthy control participants performed better than the schizophrenia patients on the MCCB, the UPSA-VIM, and all three indices from the VRFCAT, with all differences significant at $p<.001$. The magnitude of differences between the groups was very large and very similar for all outcomes, with Cohen's d=1.21 for the VRFCAT time to completion, d=1.22 for the MCCB and d=1.16 for the UPSA-VIM. For the VRFCAT total errors and progression, the group differences were smaller, but still reflected large effect sizes. One patient performed at greater than one standard deviation better than the healthy control mean on the VRFCAT at the first assessment, while 3 patients did so on the UPSA-VIM, suggesting that ceiling effects were rare for both measures.

3.2 Test-Retest Stability and practice effects

In Table 4, we present the test-retest stability and practice effects for the VRFCAT and UPSA-VIM. The UPSA-VIM had a significantly larger practice effect for both participant samples. The test-retest coefficients for the VRFCAT total time and UPSA-VIM scores were consistent, especially for patients, with errors and forced progressions having lower test-retest reliability. No VRFCAT variables evidenced any substantial changes associated with re-testing in either sample. The lack of retest differences suggests similar levels of difficulty for the 6 versions of the VRFCAT.

3.3 Correlations with other performance-based measures

Table 5 contains Pearson correlations between the VRFCAT outcome measures, the UPSA-VIM total score and the MCCB composite score for healthy controls and patients with schizophrenia. Even though the VRFCAT is a computerized measure, and thus has a different method than the UPSA-VIM and 9 of the 10 tests of the MCCB, the UPSA-VIM and VRFCAT demonstrated similar correlations with the MCCB for the healthy controls (0.74 vs. 0.68, $z=1.32$, $p=0.19$). In the patient group, the correlation between the MCCB and UPSA-VIM was larger than the $r=0.65$ usually reported in the literature and was significantly greater than the VRFCAT-MCCB correlation (0.70 vs. 0.57, $z=2.41$, $p=0.016$). All correlations between the VRFCAT measures, the MCCB composite, and the UPSA-VIM were significant ($p<.001$) in patients and controls, although correlations for VRFCAT errors and progression with the UPSA-VIM and MCCB were smaller in magnitude.

3.4 Correlations with interview-based functional and cognitive outcomes

For patients with schizophrenia, the UPSA-VIM and all three VRFCAT measures (total time, total errors, and occurrence of progression) were significantly correlated with better everyday functioning on the SLOF and lower ratings of cognitive impairment on the SCoRS (Table 6). The sum of the individual items of the SCoRS was significantly correlated with the SLOF total score (r=0.57, P<.001). Almost all SCoRS items had correlations >0.20 (P<.01) with the SLOF domain scores (See Supplemental Table 1).

3.5 Tolerability for patients

Subjects on average found the VRFCAT to be highly realistic, pleasant to take, easy to use, and found the instructions clear (Supplemental Table 2). Although all subjects rated the task highly with respect to ease of use and understandability of instructions, these ratings were higher for healthy controls.

3.6 Equivalence of forms

The next analyses examined the equivalence of the 6 forms of the VRFCAT. We focused on the total time variable because of its correlation with the MCCB and its greater sensitivity to diagnosis. Supplemental figure 1 presents differences between healthy control and patient samples on the 6 forms of the VRFCAT at both visits presented as Cohen's d. As can be seen in the figure, the effect size for separation of controls and patients was fairly consistent across all the forms and all of these differences were statistically significant (all d>0.9; all n≥47; all p≤.001), although Cohen's d was smaller at the first assessment for forms 1 and 6. These results suggest that the 6 forms of the VRFCAT, administered either first or second, are associated with substantial separation of the two groups.

Also, we examined the correlations between the UPSA-VIM and the MCCB and each form of the test at the first assessment for both the total sample and just the schizophrenia patients. As can be seen in supplemental figure 2, the correlations were quite similar for total time and UPSA-VIM scores across the different forms of the VRFCAT. Similarly, the correlations for each form of the VRFCAT and the composite MCCB score were comparable to each other and essentially identical to the UPSA-VIM correlations for every form. Analogous results were found in the patient sample, although the correlations were approximately 10 percentage points lower overall. A two-way Analysis of Variance (ANOVA) of time to completion was used to compare the different VRFCAT forms for the patients and controls at baseline. An overall difference between forms was noted ($F_{5,322}$=3.28, p=.007) and pairwise comparisons using the Tukey-Kramer adjustment for multiple comparisons revealed significant differences (p≤.05) between form 1 with the lowest LSMean T-scores (37.6±1.82) and forms 2 and 5 with the highest (45.1±1.92 and 45.7±1.87, respectively). A test for the interaction of subject group and form was nonsignificant ($F_{5,322}$=0.20, p=.96).

4. Discussion

The Virtual Reality Functional Capacity Assessment Tool (VRFCAT), a computerized test involving several functionally relevant cognitive challenges, was found to separate healthy

controls and schizophrenia patients and to manifest convergent validity with the MCCB, the SCoRS, and the UPSA-VIM, the primary outcome measures previously used in successful treatment studies of cognitive impairment in schizophrenia. The test was reported to be easily tolerable for patients, and required about 30 minutes to complete.

The VRFCAT compared favorably with the test most frequently used as a co-primary measure in schizophrenia clinical trials, the UPSA. In this study, the version of the UPSA with the best psychometric characteristics from previous studies, the UPSA-VIM, was used. Test-retest reliability of the VRFCAT was similar to the UPSA-VIM, with smaller practice effects than the UPSA-VIM. These practice effect differences are likely due to the 6 alternate forms of the VRFCAT. All 6 of these forms manifested similar separation of healthy controls and schizophrenia patients, as well as correlations with the UPSA-VIM and MCCB that were essentially identical across all 6 forms. Further, 3 patients performed 1 SD above the mean for HC on the UPSA, and 1 patient did so with the VRFCAT, suggesting that all or almost all patients would have room to improve during the course of a treatment study.

In contrast to interview-based measures of cognition, the VRFCAT is purely performance-based and does not require an informant. While interview-based measures of cognition have good psychometric characteristics (Keefe et al., 2006a; Keefe et al., 2015a; Ventura et al., 2013), the SCoRS has far weaker correlations with objective performance-based measures and is less sensitive to the effects of treatments that improve performance on the MCCB when informants are not available (Keefe et al., 2015a). Performance-based functional capacity measures such as the VRFCAT do not require informants and thus may help to reduce the burden on investigators and participants in clinical trials.

In this study, completion time was the variable most robustly associated with MCCB and UPSA-VIM scores as compared to indices of poor performance such as errors and progressions. In the VRFCAT procedure, more errors and more progressions lead to longer time for completion. A time variable as an outcome is common in neuropsychological assessments. For instance, rapid completion with concurrent error correction by testers is the key performance demand in three of the MCCB tasks: Trails A, BACS symbol coding, and verbal fluency. In fact, tests with time-based dependent variables have repeatedly been found to be the strongest correlates of composite neuropsychological performance (Keefe et al., 2006b), performance on measures of functional capacity (McClure et al., 2007), and everyday outcomes (Harvey et al., 2009). These tests have long been known to have small but systematically detectable practice effects (Goldberg et al., 2010; Keefe et al., 2008) as well as substantial importance for prediction of functioning across neuropsychiatric conditions (Jaeger et al., 2006).

There are some limitations of this research design and dataset. First, it was not possible to perform a test-retest study of individual forms (i.e., form 1 at both time 1 and 2) with the current sample size; we were more interested in exploring alternative forms for their equivalence. Second, the correlations of the SCoRS and SLOF with the UPSA-VIM and VRFCAT were relatively small. This pattern of correlations is consistent with previous literature on the relationship of performance-based with interview-based outcomes (Green et al., 2008; Green et al., 2011), whereby method variance contributes so robustly to the

strength of the relationship among measures that it outweighs the relationship among constructs. Third, we did not assess patients with acute symptoms, so the practicality of using the VRFCAT in exacerbated patients is not known. Finally, we did not attempt to assess the treatment sensitivity of the VRFCAT in this phase of its development. It is important to note that the MCCB and the UPSA-VIM were also developed without this information available at the time of their development. These measures have each demonstrated sensitivity to treatment effects with both pharmacological and cognitive remediation interventions (Bowie et al., 2012; Fisher et al., 2009; Green et al., 2008; Javitt et al., 2012; Keefe et al., 2015b; Mahableshwarkar et al., 2015). Future treatment studies using the VRFCAT as a treatment target will test its sensitivity.

In summary, an immersive virtual reality functional capacity assessment with 6 forms was found to manifest high levels of reliability and convergent validity across the forms. Patient experience with computers did not pose a barrier to completion and the psychometric characteristics of the measure meet all of the established criteria for acceptance. The potential flexibility of the computerized format is also a benefit, in that the test has the potential for being delivered remotely as well as being modified for cross-cultural usage in geographical regions where computer use, public transportation, and grocery stores are common.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Disclosures

Dr. Keefe currently or in the past 3 years has received investigator-initiated research funding support from the Department of Veteran's Affair, Feinstein Institute for Medical Research, National Institute of Mental Health, Psychogenics, Research Foundation for Mental Hygiene, Inc., and the Singapore National Medical Research Council. He currently or in the past 3 years has received honoraria, served as a consultant, or advisory board member for Abbvie, Akebia, Asubio, Avanir, AviNeuro/ChemRar, BiolineRx, Biogen Idec, BiolineRx, Biomarin, Boehringer-Ingelheim, EnVivo/FORUM, GW Pharmaceuticals, Janssen, Johnson & Johnson, Lundbeck, Merck, Minerva Neurosciences, Inc., Mitsubishi, Neuralstem, Neuronix, Novartis, NY State Office of Mental Health, Otsuka, Pfizer, Reviva, Roche, Sanofi/Aventis, Shire, Sunovion, Takeda, Targacept, and the University of Texas South West Medical Center. Dr. Keefe receives royalties from the BACS testing battery, the MATRICS Battery (BACS Symbol Coding) and the Virtual Reality Functional Capacity Assessment Tool (VRFCAT). He is also a shareholder in NeuroCog Trials, Inc. and Sengenix.

Dr. Patterson currently or in the past 3 years has received consulted with NeuroCog Trials, Inc, Abbott Labs, and Amgen, and has been funded by NIMH.

Dr. Narasimhan currently or in the past 3 years has received investigator-initiated research funding support from NIMH, NIDA, NIAAA, NIH, Eli Lilly, Janssen Pharmaceuticals, Forest Labs, and Pfizer. She currently or in the past 3 years has received honoraria, served as a consultant, or advisory board member for Eli Lilly.

Dr. Harvey has received consulting fees from Acadia, Boehringer-Ingelheim, Forum Pharma, Lundbeck, Otsuka America, Sanofi Pharma, Sunovion Pharma, and Takeda Pharma during the past year. He also received a research grant from Takeda. He also receives royalties from the MCCB and for the Brief Assessment of Cognition.

Drs. Davis, Atkins, and Vaughn report being a full-time employees of NeuroCog. Trials, Inc.

Role of Funding Source

Funding provided by the National Institute of Mental Health Grant Number 1R43MH084240-01A2 and 2R44MH084240-02.

References

Atkins AS, Stroescu I, Spagnola NB, Davis VG, Patterson TD, Narasimhan M, Harvey PD, Keefe RS. Assessment of age-related differences in functional capacity using the Virtual Reality Functional Capacity Assessment Tool (VRFCAT). J Prev Alzheimers Dis. 2015; (2):121–127. [PubMed: 26618145]

Bilder RM, Goldman RS, Robinson D, Reiter G, Bell L, Bates JA, Pappadopulos E, Willson DF, Alvir JM, Woerner MG, Geisler S, Kane JM, Lieberman JA. Neuropsychology of first-episode schizophrenia: initial characterization and clinical correlates. Am J Psychiatry. 2000; (157):549–559. [PubMed: 10739413]

Bowie CR, Harvey PD. Cognition in schizophrenia: Impairment, determinants, and. functional importance. Psychiatr Clin North Am. 2005; (28):613–633. [PubMed: 16122570]

Bowie CR, McGurk SR, Mausbach B, Patterson TL, Harvey PD. Combined cognitive remediation and functional skills training for schizophrenia: effects on cognition, functional competence, and real-world behavior. Am J Psychiatry. 2012; (169):710–718. [PubMed: 22581070]

Buchanan RW, Davis M, Goff D, Green MF, Keefe RSE, Leon A, Nuechterlein K, Laughren T, Levin R, Stover E, Fenton W, Marder S. A Summary of the FDA-NIMH-MATRICS workshop on clinical trial designs for neurocognitive drugs for schizophrenia. Schizophr Bull. 2005; (31):5–19. [PubMed: 15888422]

Buchanan RW, Keefe RSE, Umbricht D, Green MF, Laughren T, Marder SR. The FDA-NIMH-MATRICS guidelines for clinical trial design of cognitive-enhancing drugs: What do we know 5 years later? Schizophr Bull. 2011; (37):1209–1217. [PubMed: 20410237]

Davidson M, Galderisi S, Weiser M, Werbeloff N, Fleischhacker WW, Keefe RS, Boter H, Keet IP, Prelipceanu D, Rybakowski JK, Libiger J, Hummer M, Dollfus S, Lopez-Ibor JJ, Hranov LG, Gaebel W, Peuskens J, Lindefors N, Riecher-Rossler A, Kahn RS. Cognitive effects of antipsychotic drugs in first-episode schizophrenia and schizophreniform disorder: A randomized, open-label clinical trial (EUFEST). Am J Psychiatry. 2009; (166):675–682. [PubMed: 19369319]

Durand D, Strassnig M, Sabbag S, Gould F, Twamley EW, Patterson TL, Harvey PD. Factors influencing self-assessment of cognition and functioning in schizophrenia: Implications for treatment studies. Eur Neuropsychopharmacology. 2015; (25):185–191.

Fisher M, Holland C, Merzenich MM, Vinogradov S. Using neuroplasticity-based auditory training to improve verbal memory in schizophrenia. Am J Psychiatry. 2009; (166):805–811. [PubMed: 19448187]

Goldberg TE, Keefe RS, Goldman RS, Robinson DG, Harvey PD. Circumstances under which practice does not make perfect: a review of the practice effect literature in schizophrenia and its relevance to clinical treatment studies. Neuropsychopharmacology. 2010; (35):1053–1062. [PubMed: 20090669]

Green MF, Nuechterlein KH, Kern RS, Baade LE, Fenton WS, Gold JM, Keefe RSE, Mesholam-Gately R, Seidman LJ, Stover E, Marder SR. Functional co-primary measures for clinical trials in schizophrenia: results from the MATRICS psychometric and standardization study. Am J Psychiatry. 2008; (165):221–228. [PubMed: 18172017]

Green MF, Schooler NR, Kern RS, Frese FJ, Granberry W, Harvey PD, Karson CN, Peters N, Stewart M, Seidman LJ, Sonnenberg J, Stone WS, Walling D, Stover E, Marder SR. Evaluation of functionally-meaningful measures for clinical trials of cognition enhancement in schizophrenia. Am J Psychiatry. 2011; (168):400–407. [PubMed: 21285142]

Harvey PD, Silverman JM, Mohs RC, Parrella M, White L, Powchik P, Davidson M, Davis KL. Cognitive decline in late-life schizophrenia: A longitudinal study of geriatric chronically hospitalized patients. Biol Psychiatry. 1999; (45):32–40. [PubMed: 9894573]

Harvey PD, Keefe RSE, Patterson TL, Heaton RK, Bowie CR. Abbreviated neuropsychological assessments in schizophrenia: association with different outcomes measures. J Clin Neuropsychol. 2009; (31):462–471.

Jaeger J, Berns S, Uzelac S, Davis-Conway S. Neurocognitive deficits and disability in major depressive disorder. Psychiatry Res. 2006; (145):39–48. [PubMed: 17045658]

Javitt DC, Buchanan RW, Keefe RSE, Kern R, McMahon RP, Green MF, Lieberman J, Goff DC, Csernansky JG, McEvoy JP, Jarskog F, Seidman LJ, Gold JM, Kimhy D, Nolan KS, Barch DS, Ball MP, Robinson J, Marder SR. Effect of the neuroprotective peptide davunetide (AL-108) on cognition and functional capacity in schizophrenia. Schizophr Res. 2012; (136):25–31. [PubMed: 22169248]

Kay SR, Fiszbein A, Opler LA. The Positive And Negative Syndrome Scale (PANSS) for schizophrenia. Schizophr Bull. 1987; (13):261–276. [PubMed: 3616518]

Keefe RSE, Poe M, Walker TM, Kang JW, Harvey PD. The Schizophrenia Cognition Rating Scale (SCoRS): an interview-based assessment and its relationship to cognition, real-world functioning and functional capacity. Am J Psychiatry. 2006a; (163):426–432. [PubMed: 16513863]

Keefe RSE, Bilder RM, Harvey PD, Davis SM, Palmer BW, Gold JM, Meltzer HY, Green MF, Miller DD, Canive JM, Adler LW, Manschreck TC, Swartz M, Rosenheck R, Perkins DO, Walker TM, Stroup TS, McEvoy JP, Lieberman JA. Baseline neurocognitive deficits in the CATIE schizophrenia trial. Neuropsychopharmacology. 2006b; (31):2033–2046. [PubMed: 16641947]

Keefe RSE, Bilder RM, Davis SM, Harvey PD, Palmer BW, Gold JM, Meltzer HY, Green MF, Capuano G, Stroup TS, McEvoy JP, Swartz MS, Rosenheck RA, Perkins DO, Davis CE, Hsiao JK, Lieberman JA, CATIE Investigators. Neurocognitive Working Group: Neurocognitive effects of antipsychotic medications in patients with chronic schizophrenia. Arch Gen Psychiatry. 2007; (64):633–647. [PubMed: 17548746]

Keefe RSE, Malhotra AK, Meltzer H, Kane JM, Buchanan RW, Murthy A, Sovel M, Li C, Goldman R. Efficacy and safety of donepezil in patients with schizophrenia or schizoaffective disorder: significant placebo/practice effects in a 12-week, randomized, double-blind, placebo-controlled trial. Neuropsychopharmacology. 2008; (33):1217–1228. [PubMed: 17625502]

Keefe RSE, Davis VG, Spagnola NB, Hilt D, Dgetluck N, Ruse S, Patterson TD, Narasimhan M, Harvey PD. Reliability, validity and treatment sensitivity of the Schizophrenia Cognition Rating Scale. Eur Neuropsychopharmacology. 2015a; (25):176–184.

Keefe RS, Meltzer HA, Dgetluck N, Gawryl M, Koenig G, Moebius HJ, Lombardo I, Hilt DC. Randomized, double-blind, placebo-controlled study of encenicline, an α-7 nicotinic acetylcholine receptor agonist as an adjunctive treatment for cognitive impairment in schizophrenia. Neuropsychopharmacology. 2015b; (40):3053–3060. [PubMed: 26089183]

Kern RS, Nuechterlein KH, Green MF, Baade LE, Fenton WS, Gold JM, Keefe RSE, Mesholam-Gately R, Mintz J, Seidman LJ, Stover E, Marder SR. The MATRICS consensus cognitive battery; Part 2: Co-norming and standardization. Am J Psychiatry. 2008; (165):214–220. [PubMed: 18172018]

Laughren T. A regulatory perspective on psychiatric syndromes in Alzheimer's disease. Am J Geriatr Psychiatry. 2001; 9(9):340–345. [PubMed: 11739061]

Leifker FR, Bowie CR, Harvey PD. The determinants of everyday outcomes in schizophrenia: Influences of cognitive impairment, clinical symptoms, and functional capacity. Schizophr Res. 2009; (115):82–87. [PubMed: 19775869]

Leifker FR, Patterson TL, Heaton RK, Harvey PD. Validating measures of real-world outcome: the results of the VALERO Expert Survey and RAND Appropriateness Panel. Schizophr Bull. 2011; (37):334–343. [PubMed: 19525354]

Mahableshwarkar AR, Zajecka J, Jacobson W, Chen Y, Keefe RS. A randomized, placebo-controlled, active-reference, double-blind, flexible-dose study of the efficacy of vortioxetine on cognitive function in major depressive disorder. Neuropsychopharmacology. 2015; (40):2025–2037. [PubMed: 25687662]

Marder SR, Fenton W. Measurement and treatment research to improve cognition in schizophrenia: NIMH MATRICS initiative to support the development of agents for improving cognition in schizophrenia. Schizophr Res. 2004; (72):5–9. [PubMed: 15531402]

McClure MM, Bowie CR, Patterson TL, Heaton RK, Weaver C, Anderson H, Harvey PD. Correlations of functional capacity and neuropsychological performance in older patients with schizophrenia: evidence for specificity of relationships? Schizophr Res. 2007; (89):330–338. [PubMed: 16982175]

Moore DJ, Palmer BW, Patterson TL, Jeste DV. A review of performance-based measures of everyday functioning. J Psychiatr Res. 2007; (41):97–118. [PubMed: 16360706]

Nuechterlein KH, Green MF, Kern RS, Baade LE, Barch DM, Cohen JD, Essock S, Fenton WS, Frese FJ III, Gold JM, Goldberg T, Heaton RK, Keefe RS, Kraemer H, Mesholam-Gately R, Seidman LJ, Stover E, Weinberger DR, Young AS, Zalcman S, Marder SR. The MATRICS Consensus Cognitive Battery, part 1: test selection, reliability, and validity. Am J Psychiatry. 2008; (165):203–213. [PubMed: 18172019]

Rosenheck R, Leslie D, Keefe R, McEvoy J, Swartz M, Perkins D, Stroup S, Hsiao JK, Lieberman J, CATIE Study Investigators Group. Barriers to employment for people with schizophrenia. Am J Psychiatry. 2006; (163):411–417. [PubMed: 16513861]

Ruse SA, Harvey PD, Davis VG, Atkins AS, Fox KH, Keefe RS. Virtual Reality Functional Capacity Assessment In Schizophrenia: preliminary data regarding feasibility and correlations with cognitive and functional capacity performance. Schizophr Res Cogn. 2014; (1):e21–e26. [PubMed: 25083416]

Sabbag S, Twamley EM, Vella L, Heaton RK, Patterson TL, Harvey PD. Assessing everyday functioning in schizophrenia: not all informants seem equally informative. Schizophr Res. 2011; (131):250–255. [PubMed: 21620682]

Saykin AJ, Shtasel DL, Gur RE, Kester DB, Mozley LH, Stanfiniak P, Gur RC. Neuropsychological deficits in neuroleptic naive patients with first episode schizophrenia. Arch Gen Psychiatry. 1994; (51):124–131. [PubMed: 7905258]

Schneider LC, Struening EL. SLOF: A behavioral rating scale for assessing the mentally ill. Soc Work Res Abstr. 1983; (19):9–21. [PubMed: 10264257]

Seidman LJ, Giuliano AJ, Meyer EC, Addington J, Cadenhead KS, Cannon TD, McGlashan TH, Perkins DO, Tsuang MT, Walker EF, Woods SW, Bearden CE, Christensen BK, Hawkins K, Heaton R, Keefe RS, Heinssen R, Cornblatt BA, North American Prodrome Longitudinal Study (NAPLS) Group. Neuropsychology of the prodrome to psychosis in the NAPLS consortium: Relationship to family history and conversion to psychosis. Arch Gen Psychiatry. 2010; (67):578–588. [PubMed: 20530007]

Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC. The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry. 1998; (59):22–33. [PubMed: 9881538]

Velligan D, Mintz J, Maples N, Xueying L, Gajewski S, Carr H, Sierra C. A randomized trial comparing in person and electronic interventions for improving adherence to oral medications in schizophrenia. Schizophr Bull. 2012; (39):999–1007. [PubMed: 23086987]

Ventura J, Reise SP, Keefe RSE, Hurford IM, Wood RC, Bilder RM. The Cognitive Assessment Interview (CAI): Reliability and validity of a brief interview-based measure of cognition. Schizophr Bull. 2013; (39):583–591. [PubMed: 22328641]

Table 1

Objectives of the VRFCAT.

Mini Scenario		Objective
Apartment	1.	Pick up the recipe on the counter
	2.	Search for ingredients in your cabinets and refrigerator
	3.	Access your recipe and cross off the ingredients that you already have in your apartment. When you are finished crossing off the ingredients, pick up the bus schedule on the counter
	4.	Pick up the billfold on the counter
	5.	Exit the apartment and head to the bus stop
Bus to Store	6.	Wait for the correct bus to the grocery store and then board it when it arrives
	7.	Add up the exact amount of bus fare in your hand and pay for the bus
Store	8.	Select a food aisle to begin shopping
	9.	Continue shopping for the necessary food ingredients, and when finished check out
	10.	Add up the exact amount for your purchase in your hand and pay for groceries
Bus to Apartment	11.	Wait for the correct bus to your apartment and then board it when it arrives
	12.	Add up the exact amount of bus fare in your hand and pay for the bus

Table 2

Descriptive information on the participant samples

	Healthy Controls			Schizophrenia Patients				
	N	**M**	**SD**	**N**	**M**	**SD**		
							t	**p**
Age	165	42.6	13.94	158	43.6	11.85	−0.72	.47
Years of Education	165	14.7	2.41	157	12.8	1.99	7.77	<.001
Mother's Education	155	12.9	2.98	142	12.5	3.33	1.18	.24
	%	(N)		%	N		X^2	p
Male	53	(88)		56	(88)		0.18	.67
Unemployed	33	(54)		85	(135)		92.40	<.001
Comfortable with Computer	97	(160)		89	(140)		8.53	.004
Hispanic	18	(29)		19	(30)		0.11	.74
English Primary Language	95	(157)		96	(151)		0.03	.86
Race								
Caucasian	56	(92)		47	(75)		3.33	.19
African American	38	(63)		48	(76)			
Other	6	(10)		4	(7)			

Schizophr Res. Author manuscript; available in PMC 2017 August 01.

Table 3

Performance on VRFCAT, MCCB, and UPSA and Clinical Variables at Visit 1

	Healthy Controls			Schizophrenia Patients			Cohen's d
	N	M	SD	N	M	SD	
MCCB[a]	163	44.0	13.19	155	28.1	12.91	1.22
UPSA-VIM[a]	164	83.2	9.03	158	71.0	11.85	1.16
VRFCAT Total Time[b]	165	49.7	11.48	158	32.5	16.59	1.21
VRFCAT Total Errors[b]	165	49.4	11.61	158	37.7	22.40	0.66
VRFCAT Progression[b]	165	49.7	10.16	158	40.5	13.64	0.77
SCoRS Interviewer Ratings[c]	–	–	–	156	38.2	9.88	
SLOF Interviewer Total Score[d]	–	–	–	158	11.1	1.64	
PANSS Total scores	–	–		158	71.6	21.93	

Notes: All Cohen's d are significant at p<.001.

[a]Maximum total score is 100;

[b]T Score: Mean =50, SD=10 in healthy controls;

[c]Maximum score is 80, higher scores are worse;

[d]Maximum score is 15.

Table 4

Test-Retest Stability and Practice Effects for UPSA-VIM and VRFCAT Variables

	Visit 1				Visit 2							
	HC (N=165)		SZ (N=158)		HC (N=159)		SZ (N=151)		Cohen's d		ICC	
	M	SD	M	SD	M	SD	M	SD	HC	SZ	HC	SZ
VRFCAT Time	49.7	11.48	32.5	16.59	50.9	11.52	31.8	17.62	.10	−.04	.65	.81
VRFCAT Errors	49.4	11.61	37.6	22.40	49.8	12.94	36.7	22.07	.03	−.04	.54	.65
VRFCAT Progressions	49.7	10.16	40.5	13.64	50.3	10.51	40.8	13.58	.06	.03	.29	.61
UPSA-VIM	83.2	9.03	71.0	11.85	86.7	9.07	74.5	12.07	.39***	.35***	.75	.78

Note

*** p<.001

HC, healthy controls
SZ, schizophrenia patients

Schizophr Res. Author manuscript; available in PMC 2017 August 01.

Table 5

Pearson Correlations between VRFCAT Variables and MCCB and UPSA Scores in Healthy Controls (N=162) and Schizophrenia Patients (N=155)

	1	2	3	4	5
1. MCCB	–	**.74**	**.68**	**.50**	**.36**
2. UPSA-VIM	.70	–	**.58**	**.50**	**.39**
3. VRFCAT total time	.57	.56	–	**.74**	**.59**
4. VRFCAT total Errors	.39	.41	.70	–	**.72**
5. VRFCAT Total Progressions	.45	.43	.71	.65	–

Note. Healthy Controls above the diagonal. Schizophrenia patients are below. All correlations are significant at p<.001. Sample size restricted to those subjects with data available for all 5 measures.

Table 6

Correlations with Convergent Validity Variables for Patients Only

	Schizophrenia Cognition Rating Scale (SCoRS) (N = 156)	Specific Levels of Functioning (SLOF) (N = 158)
UPSA-VIM	−.24**	.25**
VRFCAT Total Time	−.29***	.22**
VRFCAT Total Errors	−.28***	.29***
VRFCAT Progression	−.29***	.17*

Note: Significance of correlations

* p<.05,

** p<.01,

*** p<.001