Schizophrenia Research 181 (2017) 100–106

Contents lists available at ScienceDirect

Schizophrenia Research

journal homepage: www.elsevier.com/locate/schres

Validation of the tablet-administered Brief Assessment of Cognition (BAC App)

Alexandra S. Atkins [a], Tina Tseng [a], Adam Vaughan [a], Elizabeth W. Twamley [b,c], Philip Harvey [d], Thomas Patterson [b], Meera Narasimhan [e], Richard S.E. Keefe [a,f,*]

[a] NeuroCog Trials, Durham, NC, USA
[b] University of California, San Diego, School of Medicine, San Diego, CA, USA
[c] Center of Excellence for Stress and Mental Health, VA San Diego Healthcare System, San Diego, CA USA
[d] University of Miami Miller School of Medicine, Miami, FL, USA
[e] University of South Carolina School of Medicine, Columbia, SC, USA
[f] Duke University Medical Center, Durham, NC, USA

ARTICLE INFO

Article history:
Received 26 April 2016
Received in revised form 3 October 2016
Accepted 5 October 2016
Available online 19 October 2016

Keywords:
Schizophrenia
Cognition
App
Cognitive test
Tablet assessment
BACS

ABSTRACT

Computerized tests benefit from automated scoring procedures and standardized administration instructions. These methods can reduce the potential for rater error. However, especially in patients with severe mental illnesses, the equivalency of traditional and tablet-based tests cannot be assumed. The Brief Assessment of Cognition in Schizophrenia (BACS) is a pen-and-paper cognitive assessment tool that has been used in hundreds of research studies and clinical trials, and has normative data available for generating age- and gender-corrected standardized scores. A tablet-based version of the BACS called the BAC App has been developed. This study compared performance on the BACS and the BAC App in patients with schizophrenia and healthy controls. Test equivalency was assessed, and the applicability of paper-based normative data was evaluated. Results demonstrated the distributions of standardized composite scores for the tablet-based BAC App and the pen-and-paper BACS were indistinguishable, and the between-methods mean differences were not statistically significant. The discrimination between patients and controls was similarly robust. The between-methods correlations for individual measures in patients were $r > 0.70$ for most subtests. When data from the Token Motor Test was omitted, the between-methods correlation of composite scores was $r = 0.88$ ($df = 48$; $p < 0.001$) in healthy controls and $r = 0.89$ ($df = 46$; $p < 0.001$) in patients, consistent with the test-retest reliability of each measure. Taken together, results indicate that the tablet-based BAC App generates results consistent with the traditional pen-and-paper BACS, and support the notion that the BAC App is appropriate for use in clinical trials and clinical practice.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Cognitive deficits in schizophrenia can be severe and devastating to patients and their families. People with schizophrenia underperform controls by an estimated one and a half to two standard deviations on multiple aspects of cognitive functioning (Nuechterlein et al., 2008). Cognitive impairment in schizophrenia is an important unmet medical need, and is the target of numerous treatment efforts (Keefe et al., 2011a,b, 2013; Wykes et al., 2011; Goff et al., 2011).

A variety of methods have been used to measure cognition in patients with schizophrenia in clinical practice and clinical trials. Standard neuropsychological tests have been used for decades to measure cognition in patients with schizophrenia. These standard clinical measures have the benefit of proven methodologies and patient-rater interaction, which can be flexible based upon the needs of the patient and the clinical situation. In schizophrenia research, these measures have demonstrated sensitivity to impairments and consistent correlations with functional outcomes, which has been one of the key driving forces of interest in cognition as a treatment target.

Compared to traditional pen-and-paper neurocognitive assessments, computerized tests carry the potential to reduce rater error by leveraging automated scoring algorithms and standardized administrator instructions. However, especially in patients with severe mental illnesses and neurologic disorders, computerized tests can be difficult to implement due to the rigidity of their administration procedures and reduced adaptability to individual patient needs and understanding. These limitations have sometimes resulted in increased rates of missing data for computerized tests relative to standard paper methods (Silver et al., 2006; Harvey et al., 2013; Keefe et al., 2007a,b; Iverson et al., 2009). Further, because computerized tests tend to be more abstract

* Corresponding author at: Box 3270, Duke University Medical Center, Durham, NC 27710, USA.
E-mail address: Richard.keefe@duke.edu (R.S.E. Keefe).

http://dx.doi.org/10.1016/j.schres.2016.10.010
0920-9964/© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

than paper tests with physical materials, severely impaired populations may struggle to understand test requirements (Iverson et al., 2009).

One potential solution for impaired populations that may combine the strengths of both traditional pen-and-paper and computerized testing platforms is a streamlined tablet-based interface that combines the flexibility and human interaction of standard methods with the consistency and automaticity of computerized methods. Tablet-based versions of other standard cognitive assessment instruments such as the Wechsler Adult and Child Intelligence Scales (Wechsler, 2003, 2008) have been developed and are in use in clinical practice.

The Brief Assessment of Cognition in Schizophrenia (Keefe et al., 2004) was developed as a quick and effective tool for the assessment of cognitive change in schizophrenia. The BACS comprises six tasks for the evaluation of 4 cognitive domains identified as important for clinical trials in schizophrenia by the MATRICS Neurocognition Committee (Nuechterlein et al., 2004), including verbal memory, working memory, processing speed, and reasoning/problem solving. The composite score has high test-retest reliability in patients with schizophrenia and healthy controls (ICCs > 0.80), has been shown to be as sensitive to the cognitive deficits of schizophrenia as a standard 2.5-hour battery (Keefe et al., 2004) and is highly correlated ($r = 0.84$, $p < 0.001$) with the composite score derived from the CATIE Neurocognitive Test Battery (Keefe et al., 2007a,b; Hill et al., 2008). The BACS has clear functional relevance, as evidenced by strong correlations between the composite score and functional measures such as independent living skills ($r = 0.45$), performance-based assessments of performance of everyday living skills ($r = 0.56$), and interview-based assessments of cognition in patients with schizophrenia ($r = 0.48$; Keefe et al., 2006). It has also demonstrated sensitivity to treatment (Bowie et al., 2012; Geffen et al., 2012). The original BACS has been adapted for use for indications beyond schizophrenia (Keefe et al., 2014), and in this context is referred to simply as the BAC (Brief Assessment of Cognition). Normative data based upon 400 healthy participants demographically matched to the US Census (Keefe et al., 2008) allow for calculation of standardized domain scores and overall composite scores with correction for age and gender.

The BAC App was developed to allow tablet-based delivery and scoring of all subtests included in the original pen and Paper BAC/BACS instrument, which we will refer to here as the Paper BAC. The test battery can be administered in full or with individual tests to yield subtest and total scores with normative results. Like the original instrument, the BAC App assesses multiple cognitive domains. By ensuring standardized administration of instructions and test stimuli, the BAC App allows for reduction in error variance due to rater inconsistencies, and provides automated scoring. Results can be immediately reviewed on the tablet device, or transferred to a central data repository for later analysis.

The primary aim of the present validation study was to assess the validity of the BAC App as a measure of cognitive function by comparing performance and psychometric characteristics of the BAC App with the original Paper BAC in the same group of participants. In order to assess test equivalency and determine the sensitivity of the BAC App to cognitive deficits in schizophrenia, healthy controls and patients with schizophrenia completed the BAC App and Paper BAC within a single visit.

2. Methods

2.1. Participants

Participants included 48 patients (23 females) with schizophrenia (DSM-IV-TR Criteria) and 50 healthy controls (25 females) recruited from three academic sites including the University of California-San Diego, the University of Miami - Miller School of Medicine, and the University of South Carolina. All participants were screened for alcohol and substance abuse using modules J and K of the MINI International Neuropsychiatric Interview. Any subject meeting criteria for current alcohol or substance abuse was excluded. All subjects provided signed informed consent prior to completion of study-related activities.

Detailed demographic information is displayed in Table 1. To ensure adequate sampling to support validation of the BAC App across age and sex demographics, enrollment was stratified to include balanced representation of men and women in each of 6 age groups: 18–29, 30–39, 40–49, 50–59, 60–69, 70+. In addition, study investigators made a concerted effort to match patients and controls on age, sex and race. As indicated in Table 2, groups were well matched on each of these variables and on maternal education. Patients and controls differed significantly with respect to years of education attained ($p < 0.001$).

2.2. Design and procedure

All subjects completed the BAC App and the Paper BAC assessments at a single study visit, with order of administration counterbalanced across subjects. For subtests with multiple forms (Verbal Memory, Tower of London, Symbol Coding), alternate versions were utilized during administration of the tablet and paper-based tests. As such, no word list memoranda, Tower of London item or Symbol Coding configuration was completed more than once for any participant.

Following informed consent procedures, participants completed a brief demographic and medical history questionnaire. Participants were screened for current alcohol and substance abuse to ensure eligibility prior to testing. The BAC App and Paper BAC tests were then administered by trained raters. Following completion of both assessments, subjects completed a brief questionnaire to provide subjective ratings of their experiences with the BAC App. Finally, all subjects were administered two measures of functional capacity, including the Virtual Reality Functional Capacity Assessment Tool (Ruse et al., 2014; Atkins et al., 2015) and the UCSD Performance-based Skills Assessment (Patterson et al., 2001) not discussed here. The study visit lasted approximately 3 h, including screening. Participants received a $50 gift card as compensation for their time and effort.

2.2.1. BAC App method

Detailed descriptions of the Paper BAC tests can be found in previous publications (e.g. Keefe et al., 2004, 2006, 2008). The BAC app is a tablet-based version of the instrument developed for use on an iPad© tablet. Subtests within the BAC App can be administered in sequence as a full battery, as presented here, or the App can be configured to allow for administration of a smaller selection of tests and/or customized test order. In order to facilitate repeated testing and reduce practice effects, alternate versions are available for several BAC App subtests, including Verbal Memory (7 versions), Symbol Coding (8 versions), and Tower of London (8 versions).

All tests administered within the BAC App are completed under the supervision of a trained rater. Instructions for each subtest are presented by a female narrator within the App, ensuring consistent and accurate administration. The rater can initiate repetition of instructions through the touch screen to ensure adequate understanding by the subject prior to task initiation. The electronic tablet device is passed between the rater and the subject during testing.

Table 1
Sample demographics.

Parameter	Patients ($n = 48$)	Healthy controls ($n = 50$)	p value
Age	46.04 ± 13.18	48.26 ± 14.52	ns
Education	12.63 ± 1.65	15.12 ± 2.55	<0.001
Maternal education	12.47 ± 3.51	13.43 ± 2.64	ns
Sex (% male)	52%	50%	ns
Race (% white)	74%	79%	ns

All table entries are mean ± SD or %. Continuous variables were compared by independent *t*-test and the categorical variables were compared by Fisher's exact test. All p-values are non-significant ($\alpha > 0.1$) except for Education ($p < 0.001$).

Table 2
Mean performance of BAC app and Paper BACS in healthy controls and patients with schizophrenia.

	Schizophrenia				Healthy controls					
	BAC app (mean ± SD)	Paper BAC (mean ± SD)	Cohen's d	p-Value$_{SCZ}$	BAC app (mean ± SD)	Paper BAC (mean ± SD)	Cohen's d	p-Value$_{HC}$	d_{App}	d_{Paper}
Verbal Memory	32.47 ± 10.39	33.02 ± 9.49	−0.06	0.74	42.68 ± 10.59	43.22 ± 10.73	−0.06	0.80	0.97	1.02
Digit Sequencing	15.84 ± 4.03	17.02 ± 3.86	−0.30	0.15	19.52 ± 4.48	20.24 ± 4.00	−0.17	0.40	0.87	0.83
Verbal Fluency	42.09 ± 13.45	42.94 ± 13.15	−0.06	0.66	57.02 ± 12.63	53.74 ± 12.92	0.26	0.34	1.16	0.84
Symbol Coding	32.65 ± 14.01	43.02 ± 12.60	−0.79	0.0002	46.2 ± 15.99	55.82 ± 14.75	−0.63	0.002	0.91	0.94
Token Motor Task	45.42 ± 27.02	46.79 ± 13.06	−0.07	0.75	72.56 ± 31.49	60.28 ± 14.22	0.51	0.01	0.93	1.00
Tower of London	13.62 ± 4.08	13.54 ± 5.12	0.02	0.87	16.44 ± 4.27	16.14 ± 4.20	0.07	0.72	0.63	0.56
Composite	25.96 ± 14.72	29.90 ± 14.40	−0.27	0.19	47.00 ± 16.98	46.76 ± 12.95	0.02	0.94	1.34	1.25
Modified composite	31.75 ± 13.36	33.35 ± 13.75	−0.12	0.56	48.22 ± 13.72	48.38 ± 12.76	−0.01	0.95	1.23	1.15

p-Value$_{HC}$ = significance value for BAC App vs Paper BAC in healthy controls, by *t*-test (two sample independent); p-Value$_{SCZ}$ = significance value for BAC App vs Paper BAC in patients with schizophrenia, by *t*-test.

The BAC App supports audio capture for verbal subject responses. The audio recording feature is enabled on all tests requiring verbal responses, including Verbal Memory, Digit Sequencing, Semantic and Letter Fluency. In order to ensure accurate scoring of these tests, the rater is given the opportunity to review the audio files and confirm scoring at the end of test administration.

Following is a description of the six subtests of the BAC App:

Verbal Memory (*Verbal Memory & Learning Domain*): Subjects hear a list of 15 words to remember. Words are presented by the App in at a standard rate, eliminating the effects of rater variability. Alternative forms available: 7.
Outcome measure: Total number of words recalled across 5 learning trials.
Completion time: 7 min.
Digit Sequencing (*Working Memory Domain*): Subjects are presented with randomly ordered auditory clusters of numbers (e.g. 936) with steadily increasing trial length. As in Verbal Memory, items are presented by the App at a standard rate, eliminating the effects of rater variability. Subjects are asked to report the numbers in order, from lowest to highest.
Outcome measure: Number of trials with all items in the correct order.
Completion time: 5 min.
Token Motor Task (*Motor Function*): Subjects are presented with a virtual bowl and a supply of virtual tokens and asked to swipe a token from each side of the tablet with the index finger from each hand simultaneously and release them into the center container as rapidly as possible for 60 s.
Outcome measure: Number of tokens correctly dragged into the container.
Completion time: 3 min.
Semantic Fluency & Letter Fluency Tasks (*Verbal Fluency Domain*): During Semantic Fluency, subjects are given 60 s to generate as many words as possible within the category 'animals'. During Letter Fluency, subjects are asked to generate as many words as possible beginning with a given letter. Subjects are administered two trials using letters F and S.
Outcome measures: Total words generated for each fluency task. Total scores from both tasks are combined to produce the Vernal Fluency domain score.
Completion time: 5 min.
Symbol Coding (*Speed of Processing*): Subjects assign numbers to non-meaningful symbols with the use of a key that is provided. Numbers are entered on the digital keypad and appear in the location below the corresponding symbol. Following instructions and practice, subjects are given 90 s to complete as many items as possible. Alternative forms available: 8.
Outcome measure: Number items completed correctly within the 90 second test period.
Completion time: 3 min.
Tower of London (*Executive Functions/Reasoning and Problem Solving*): Subjects are shown two images presented on opposite sides of the tablet screen. Each image shows a different configuration of 3 colored balls arranged on 3 pegs. The subject is required to accurately determine the total number of times the balls in one picture would have to be moved in order to make the arrangement of balls identical to that of the other, opposing picture, while employing the standard rules employed in tower tests (balls are moved one at a time and balls on top of other balls must be moved first). Alternative forms available: 8.
Outcome measure: Number of correct responses.
Time: 7 min.

2.3. Data analyses

2.3.1. Composite scores

The primary outcome measure from the both the BAC App and Paper BAC is a composite T score that averages the standardized scaled scores from each of the six tests. The T-score is a metric for expressing standard scores and has a mean of 50 and an SD of 10. Composite scores for the Paper BAC and BAC App were computed using published normative values for the paper version of the test (Keefe et al., 2008), with adjustments for age and gender applied to T-scores for each subtest. The Composite score for subjects with missing subtests was imputed using the average of the remaining subtests.

2.3.2. Performance differences

Differences in performance on the BAC App and Paper BAC were evaluated using the *t*-test and Cohen's d. Between-group comparisons (*t*-test and Cohen's d) assessed the effectiveness of the BAC App and Paper BAC in distinguishing between patients and healthy controls. Pearson correlations between the BAC App and Paper BAC domains were calculated to determine correspondence between performance on the Paper BAC and BAC App versions of each test, as well as the summary composite scores derived from paper and iPad based assessments.

After comparing the T-scores for each subtest and the composite between the BAC App and Paper BAC, a modified composite was calculated (see the Results and Discussion sections). The modified composite score included a 10-point addition to the BAC app symbol coding and removal of the token motor subtest. R-squared was calculated to measure the correlation between the BAC app and Paper BAC modified composite scores.

3. Results

3.1. Completion rate

One patient was unable to complete the BAC App Tower of London assessment due to a lack of understanding of the task. For three healthy controls and one patient, the BAC App Verbal Fluency subtest scores were missing due to an administration error. BAC App Verbal Memory data was missing for one patient due to an administration error. For all cases with missing data, composite scores were imputed using the average of the remaining tests.

3.2. Performance differences

Table 2 displays means and standard deviations for healthy controls and schizophrenia patients on each subtest in the BAC App and Paper BAC. In both groups, the distributions of standardized composite scores for the BAC App and Paper BAC were indistinguishable, and the between-methods mean differences were not statistically significant (Table 2). Significant differences were found between the Paper BAC and the BAC App in both groups for the symbol-coding test and for the token motor test in the controls only, with the BAC app leading to poorer performance on symbol coding and better performance on the token motor test. For all other tests, the differences between methods were small and not statistically significant.

Effect sizes (Cohen's d) were also calculated between healthy controls and patients with schizophrenia within BAC app and Paper BACS. The discrimination between patients and controls was similar for composite scores from the BAC App ($d = 1.34$) and the Paper BAC ($d = 1.25$). Each test in the BAC App demonstrated sensitivity to cognitive deficits in the schizophrenia group, with patients significantly underperforming controls (see Fig. 1).

3.3. Correlations between BAC app and Paper BAC measures

The correlations between the BAC app and Paper BAC measures for healthy controls and patients with schizophrenia are presented in Table 3. In healthy controls, correlations between BAC App and Paper BAC measures were greater than $r = 0.70$ ($p < 0.01$) for Verbal Memory, Digit Sequencing, Verbal Fluency, Symbol Coding and Tower of London. The Paper BAC and BAC App versions of the Token Motor Task were more modestly correlated ($r = 0.47$, $p < 0.01$), reflecting the considerable differences between the finger swipe action required by the BAC App and the manual manipulation of tokens in the original Paper BAC. The magnitude of correlations between the BAC App and the Paper BAC was similar in the patient group: Verbal Memory, Digit Sequencing, and Verbal Fluency were correlated at $r = 0.70$ ($p < 0.01$) or greater, though the correlation for the Tower of London task was slightly lower at $r = 0.61$ ($p < 0.01$). As with healthy controls, the correlation was lowest for the Token Motor Task ($r = 0.43$, $p < 0.01$) suggesting both populations were sensitive to differences in task demands between the Paper BAC and BAC App versions of this test.

3.4. Modification of BAC App and Paper BAC composite scores

Composite scores for the Paper BAC and BAC App were computed using published normative values for the paper version of the test (Keefe et al., 2008). Using this method, the correlation between BAC App and Paper BAC composite scores for healthy controls was $r = 0.86$ and for patients $r = 0.77$ (Fig. 2a). Observed differences between the Paper BAC and BAC App on the symbol coding and token motor subtests (see Table 2) prompted further consideration of the equivalency of the paper- and tablet-based version of these measures. For the symbol coding subtest, a 10-point difference was observed between the two methods in both patients, though the standard deviations for the measures were equivalent in both groups. For the token motor test, differences were more substantial and were marked by low correlations in both patients and controls ($p < 0.50$) as well as increased variability in the BAC App version of the measure. In order to address these issues, a modified composite score for the BAC App was produced to improve test equivalency in the absence of a full set of normative data for the BAC App. Modification included a raw score adjustment of 10 points (+10) to BAC App raw scores for the symbol coding task and removal of token motor subtest from both BAC App and Paper BAC composite calculations. The correlation between modified composite scores for the BAC App and Paper BAC was $r = 0.88$ ($p < 0.001$) in healthy controls and $r = 0.89$ ($p < 0.001$) in patients (Fig. 2b), which is consistent with the test-retest reliability of the Paper BAC measure (Keefe et al., 2004). Fig. 3 presents the overlap histograms of the modified composite scores. Distributions of modified composite scores were nearly identical for the BAC App and the Paper BAC in both healthy controls (Fig. 3a) and patients with schizophrenia (Fig. 3b). Modified composite score for both the BAC App and Paper BAC demonstrated strong discrimination between patients and controls (Fig. 3c and d, respectively), indicating the modified composite is effective in increasing test equivalency while preserving sensitivity to cognitive impairments in the patient group.

3.5. Testing order

In order to evaluate the influence of testing order on observed results, BAC App modified composite scores and Paper BAC composite

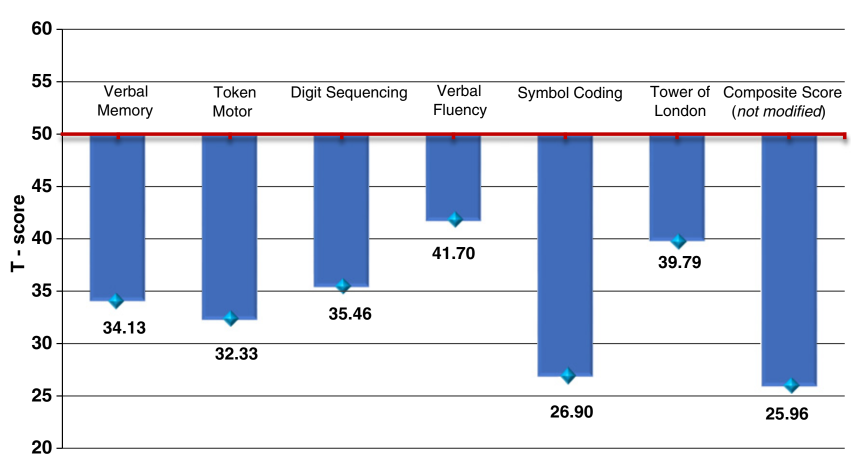

Fig. 1. Sensitivity of BAC App measures to cognitive impairment in the schizophrenia group. For all BAC App subtests, patients performed approximately 1–2 SD below healthy controls. The degree of these deficits is comparable to deficits observed using the traditional pen-and-paper BACS (e.g. Keefe et al., 2004).

Table 3
Pearson correlations between BAC App and Paper BAC tests in healthy controls and schizophrenia patients.

Paper BAC	Healthy controls						Schizophrenia patients					
	BAC App						BAC App					
	VM	DS	VF	SC	TM	TL	VM	DS	VF	SC	TM	TL
VM	**0.81****	0.39**	0.48**	0.55**	0.60**	0.41**	**0.78****	0.56**	0.52**	0.46**	0.11	0.58**
DS	0.43**	**0.83****	0.57**	0.43**	0.39**	0.46**	0.56**	**0.80****	0.49**	0.33*	0.24	0.33*
VF	0.48**	0.66**	**0.78****	0.42**	0.46**	0.48**	0.47**	0.52**	**0.93****	0.41**	0.25	0.46**
SC	0.36*	0.49**	0.37*	**0.74****	0.68**	0.46**	0.49**	0.41**	0.57**	**0.72****	0.48**	0.49**
TM	0.28*	0.36**	0.29	0.47**	**0.47****	0.32*	0.21	0.29*	0.48**	0.48**	**0.43****	0.36*
TL	0.36*	0.37**	0.39**	0.37**	0.43**	**0.74****	0.51**	0.36*	0.31*	0.26	0.01	**0.61****

BAC measures: VM = Verbal Memory; DS = Digit Sequencing; VF = Verbal Fluency; SC = Symbol Coding; TM = Token Motor Task, TL = Tower of London.
Bolded items represent correlations between Paper BAC and BAC App versions of each test.
* $p < 0.05$.
** $p < 0.01$.

scores were entered into a mixed model for repeated measures (MMRM) with Test Modality (tablet vs. paper) and with Test Order (BAC App first or second) as a between subject variable. Results indicated no main effect for the order in which tests were given ($p = 0.62$), though a Test Modality by Test Order interaction demonstrated that, as expected, subject's performed better on the second test administered ($F = 22.75$, $p < 0.001$). As such, subjects who completed the BAC App first demonstrated improvement on the Paper BAC (administered second) and subjects who completed the Paper BAC first demonstrated improvement on the BAC App.

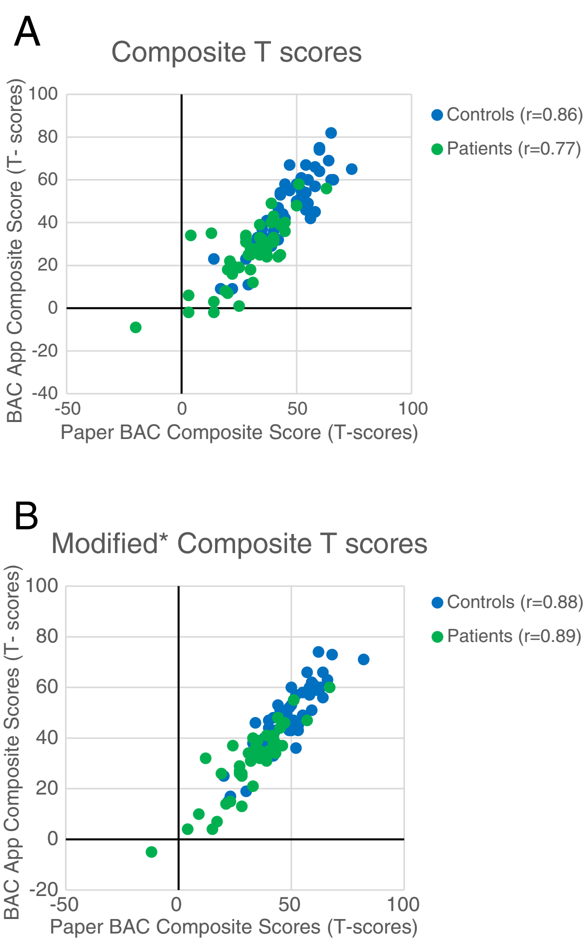

Fig. 2. Scatterplot of BAC App and Paper BAC unmodified (panel A) and modified* (panel B) composite scores for schizophrenia patients and controls. *Note: Modified composite scores incorporated a 10-point adjustment to the BAC App Symbol Coding raw score prior to standardization, and elimination of the Token Motor Test from composite calculations for both the BAC App and Paper BAC.

3.6. Correlations with the Virtual Reality Functional Capacity Assessment Tool (VRFCAT)

The correlations between the BAC App and the VRFCAT, a computer-based measure of functional capacity, were robust for healthy controls ($r = 0.69$) and for patients ($r = 0.61$). These correlations were slightly larger than those found between the VRFCAT and the Paper BAC in healthy controls ($r = 0.62$) and patients ($r = 0.46$).

4. Discussion

The primary aim of the present validation study was to assess the validity of the BAC App as a measure of cognitive function by examining performance on the BAC App and comparing it to performance on the original Paper BAC. The BAC App allows tablet-based delivery and scoring of 5 key subtests for the assessment of cognitive impairment included in the original pen and Paper BAC. By ensuring standardized administration of instructions and test stimuli, the BAC App allows for reduction in error variance due to rater inconsistencies, and provides automated scoring, yet maintains the participant-rater interactions inherent in human testing. The data from this study suggest that the BAC App has very similar overall psychometric characteristics to the traditional Paper BAC. Specifically, the BAC App had nearly identical sensitivity to the overall cognitive deficits measured by a composite score comparing patients with schizophrenia to healthy controls. The completion rate with the BAC App was very high, with very few missing data points, in contrast to computerized technology that historically has led to a significantly increased missing data rate (Silver et al., 2006; Harvey et al., 2013; Keefe et al., 2007a; Iverson et al., 2009). Finally, the BAC App demonstrated strong correlations with scores on the VRFCAT, a measure of functional capacity, suggesting that the BAC App, like the BAC, measures aspects of cognition that are relevant for the performance of everyday functional tasks.

Most cognitive tests included in the BAC App demonstrated very high correlations with the tests from the Paper BAC. These tests included Verbal Memory, Digit Sequencing, Tower of London, and Verbal Fluency. The consistency of results for these tests is natural and expected in light of minimal change in testing methodology between the Paper BAC and the BAC App. In contrast, the test method for the symbol coding test varies considerably between the Paper BAC and the BAC App, the former requiring written responses and the latter requiring the use of a digital keypad. Despite these variations, the decrement in performance from the Paper BAC to the BAC App was fairly linear with correlations >0.70 in both groups and a 10-point difference between the two methods in each group. Thus, a 10 point correction to the BAC App score enables use of the original BAC norms with minimal validity concerns. The methodology for the BAC App Token Motor Test also varies considerably from that employed by the Paper BAC. In the Paper BAC, subjects are required to manipulate real plastic tokens by lifting them off of a table and

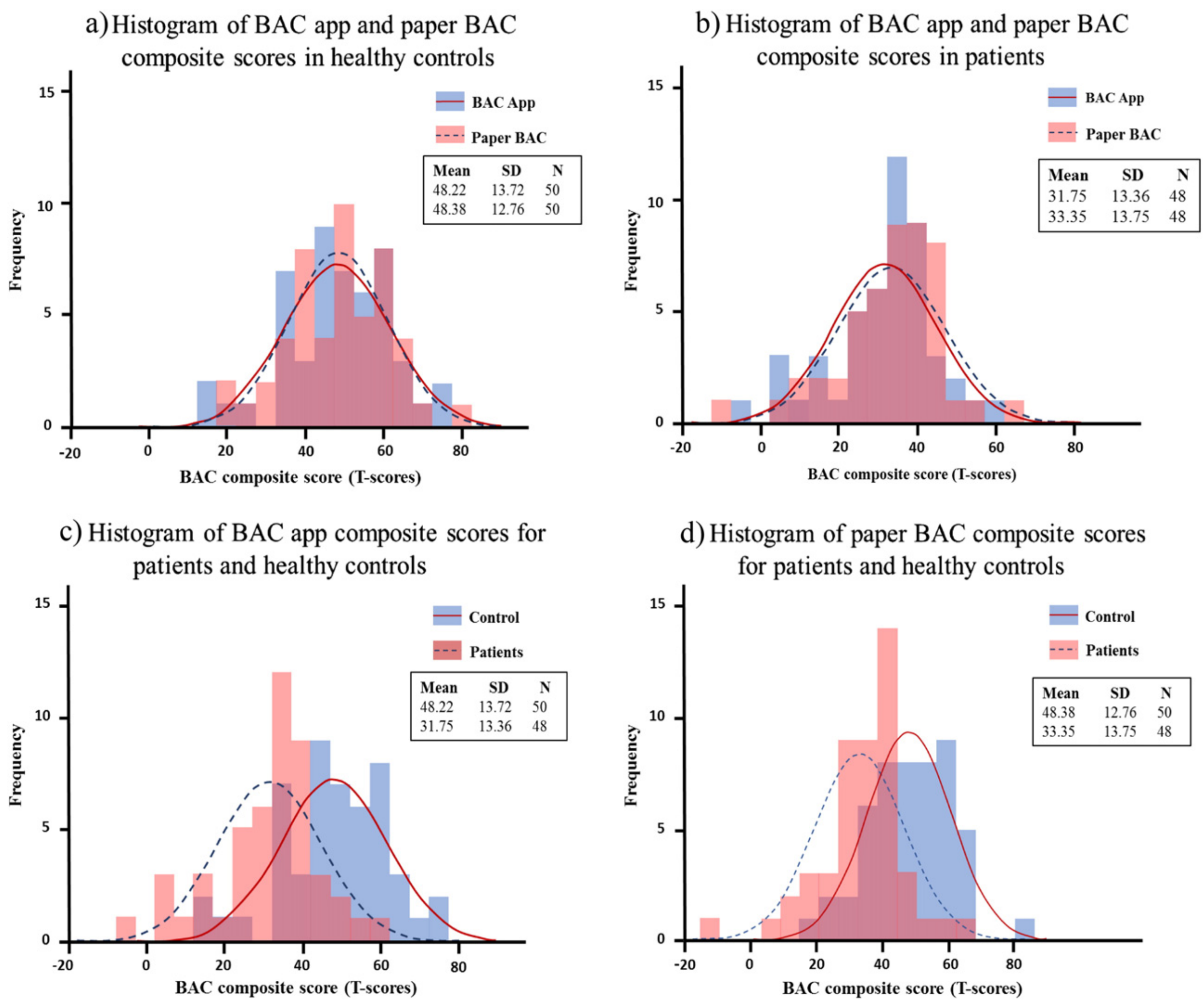

Fig. 3. Overlapping distributions of BAC app and Paper BAC composite scores for patients and healthy controls are displayed in panels a and b, respectively. Distributions for both participant group on the BAC app and Paper BAC are displayed in panels c and d, indicating the sensitivity of both instruments to group differences.

placing them into a container. By comparison, the BAC App requires subjects to swipe virtual 'digital' tokens and release them into a virtual bowl. For this test, differences in methodology likely explain the lower correlations observed between the Paper BAC and BAC App in both healthy controls and patients with schizophrenia ($r < 0.50$), as well as larger standard deviations for the BAC App version, and mean differences between the methodologies in the patient group. It is also likely that prior experience with tablet applications may be more predictive of performance on this test than any of the other tests. Future studies will need to establish normative performance for the BAC App version of the Token Motor Test, since the current norms are not applicable. While the ability to swipe a screen is an important functional skill, useful for both tablet devices and mobile phones, variability in this skill leads to divergent performance between the BAC App and the Paper BAC.

The creation of a composite score for the BAC App using previously published norms for the Paper BAC reflects very similar performance if a 10-point correction is made for the symbol coding test and the Token Motor Test is not included. This 5-test composite score was found to be correlated with the Paper BAC composite score with values of 0.89 for controls and 0.88 for patients, which is almost identical to the test-retest reliability of the Paper BAC. Thus, this score should be a useful psychometric index of cognitive functioning in patients with schizophrenia, based upon age- and sex-corrected values from the healthy population. Further, the strong correlation of 0.61 with a measure of functional capacity, the VRFCAT, in patients with schizophrenia is consistent with previous reports on the correlation between other tests batteries, such as the MCCB, with traditional measures of functional capacity such as the UCSD Performance-based Skills Assessment (Leifker et al., 2010).

There are a few limitations to this study. First, the sample size for this study was relatively small for a validation study. Although the results were very clear, and larger samples would be unlikely to reach alternate conclusions, larger samples would provide better reflection of the overall population, establish the test-retest reliability of the BAC App, and enable an investigation of the different demographic factors that may affect performance on a tablet-based cognitive test battery. Such studies are underway. Second, although we expect that the standardized administration provided by the BAC App may increase data quality in clinical trials employing raters of varying expertise, the present investigation cannot speak directly to this point. Finally, although every effort was made to ensure impartial data collection and analysis, funding for the current study was provided by a company with financial interest in the BAC and BAC App. Replication by independent research groups will be important to confirm the validity and usefulness of the BAC App.

Funding body

Funding provided by NeuroCog Trials, Durham, NC.

Contributors

Alexandra S. Atkins supervised development of the BAC App, contributed to the experimental design, analysis and interpretation of findings, and drafted several portions of the manuscript. Vicki Davis contributed to data analysis and interpretation of findings. Tina Tseng conducted statistical analyses and contributed to the creation figures. Adam Vaughan provided administrative oversight of data collection and manuscript review. Philip Harvey directed collection of data at University of Miami Miller School of Medicine, contributed to the interpretation of findings, and provided scientific and editorial review of the manuscript. Tom Patterson directed collection of data at University of California, San Diego and provided scientific and editorial review of the manuscript. Meera Narasimhan directed collection of data at University of South Carolina. Richard S. E. Keefe directed and supervised development of the BAC App, contributed to the experimental design, analysis and interpretation of findings, and drafted portions of the manuscript.

Conflict of interest

Dr. Atkins is a full-time employee of NeuroCog Trials and has received investigator-initiated support from the National Institute of Mental Health currently or in the past 3 years.

Drs. Davis, Tseng and Vaughn are employees of NeuroCog Trials.

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.

Dr. Patterson currently or in the past 3 years has received investigator-initiated support from the National Institute of Mental Health and served as a consultant for NeuroCog Trials, Abbott Labs, and Amgen.

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. 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 and the MATRICS Battery (BACS Symbol Coding). He is also a shareholder in NeuroCog Trials and Sengenix.

Acknowledgements

We thank our study participants and their families for their time and commitment.

We also thank research staff at the University of Miami Miller School of Medicine, the University of California, San Diego, and the University of South Carolina School of Medicine for supporting recruitment and data collection.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.schres.2016.10.010.

References

Atkins, A.S., Stroescu, I., Spagnola, N.B., Davis, V.G., Patterson, T.D., Narasimhan, M., Harvey, P.D., Keefe, R.S.E., 2015. Assessment of age-related differences in functional capacity using the Virtual Reality Functional Capacity Assessment Tool (VRFCAT). J. Prev. Alzheimers Dis. 2 (2), 121–127.

Bowie, C.R., McGurk, S.R., Mausbach, B., Patterson, T.L., Harvey, P.D., 2012. Combined cognitive remediation and functional skills training for schizophrenia: effects on cognition, functional competence, and real-world behavior. Am. J. Psychiatry 169 (7), 710–718.

Geffen, Y., Keefe, R., Rabinowitz, J., Anand, R., Davidson, M., 2012. BL-1020, a new γ-aminobutyric acid-enhanced antipsychotic: results of six-week, randomized, double-blind, controlled, efficacy and safety study. J. Clin. Psychiatry 73 (9), 1168–1174.

Goff, D.C., Hill, M., Barch, D., 2011. The treatment of cognitive impairment in schizophrenia. Pharmacol. Biochem. Behav. 99 (2), 245–253.

Harvey, P.D., Siu, C.O., Hsu, J., Cucchiaro, J., Maruff, P., Loebel, A., 2013. Effect of lurasidone on neurocognitive performance in patients with schizophrenia: a short-term placebo- and active-controlled study followed by a 6-month double-blind extension. Eur. Neuropsychopharmacol. 23 (11), 1373–1382.

Hill, S.K., Sweeney, J.A., Hamer, R.M., Keefe, R.S.E., Perkins, D.O., Gu, H., McEvoy, J.P., Lieberman, J.A., 2008. Efficiency of the CATIE and BACS neuropsychological batteries in assessing cognitive effects of antipsychotic treatments in schizophrenia. J. Int. Neuropsychol. Soc. 14 (2), 209–221.

Iverson, G.L., Brooks, B.L., Ashton, V.L., Johnson, L.G., Gualtieri, C.T., 2009. Does familiarity with computers affect computerized neuropsychological test performance? J. Clin. Exp. Neuropsychol. 31 (5), 594–604.

Keefe, R.S.E., Goldberg, T.E., Harvey, P.D., Gold, J.M., Poe, M., Coughenour, L., 2004. The Brief Assessment of Cognition in Schizophrenia: reliability, sensitivity, and comparison with a standard neurocognitive battery. Schizophr. Res. 68 (2–3), 283–297.

Keefe, R.S.E., Poe, M., Walker, T.M., Harvey, P.D., 2006. The relationship of the Brief Assessment of Cognition in Schizophrenia (BACS) to functional capacity and real-world functional outcome. J. Clin. Exp. Neuropsychol. 28 (2), 260–269.

Keefe, R.S.E., Bilder, R.M., Davis, S.M., Harvey, P.D., Palmer, B.W., Gold, J.M., Meltzer, H.Y., Green, M.F., Capuano, G., Stroup, T.S., McEvoy, J.P., Swartz, M.S., Rosenheck, R.A., Perkins, D.O., Davis, C.E., Hsiao, J.K., Lieberman, J.A., for the CATIE Investigators and the Neurocognitive Working Group, 2007a. Neurocognitive effects of antipsychotic medications in patients with chronic schizophrenia in the CATIE trial. Arch. Gen. Psychiatry 64 (6), 633–647.

Keefe, R.S.E., Sweeney, J.A., Gu, H., Hamer, R.M., Perkins, D.O., McEvoy, J.P., Lieberman, J.A., 2007b. Effects of olanzapine, quetiapine, and risperidone on neurocognitive function in early psychosis: a randomized, double-blind 52 week comparison. Am. J. Psychiatry 164 (7), 1061–1071.

Keefe, R.S.E., Harvey, P.D., Goldberg, T.E., Gold, J.M., Walker, T.M., Kennel, C., Hawkins, K., 2008. Norms and standardization of the Brief Assessment of Cognition in Schizophrenia (BACS). Schizophr. Res. 102 (1–3), 108–115.

Keefe, R.S.E., Fox, K.H., Harvey, P.D., Cucchiaro, J., Siu, C., Loebel, A., 2011a. Characteristics of the MATRICS consensus cognitive battery in a 29-site antipsychotic schizophrenia clinical trial. Schizophr. Res. 125 (2–3), 161–168.

Keefe, R.S.E., Vinogradov, S., Medalia, A., Silverstein, S.M., Bell, M.D., Dickinson, D., Ventura, J., Marder, S.R., Stroup, T.S., 2011b. Report from the working group conference on multi-site trial design for cognitive remediation in schizophrenia. Schizophr. Bull. 37 (5), 1057–1065.

Keefe, R.S., Buchanan, R.W., Marder, S.R., Schooler, N.R., Dugar, A., Zivkov, M., Stewart, M., 2013. Clinical trials of potential cognitive-enhancing drugs in schizophrenia: what have we learned so far? Schizophr. Bull. 39 (2), 417–435.

Keefe, R.S.E., Fox, K.H., Davis, V.G., Kennel, C., Walker, T.M., Burdick, K.E., Harvey, P.D., 2014. The Brief Assessment of Cognition in Affective Disorders (BAC-A): performance of patients with bipolar depression and healthy controls. J. Affect. Disord. 166, 86–92.

Leifker, F.R., Patterson, T.L., Bowie, C.R., Mausbach, B.T., Harvey, P.D., 2010. Psychometric properties of performance-based measurements of functional capacity: test-retest reliability, practice effects, and potential sensitivity to change. Schizophr. Res. 119 (1–3), 246–252.

Nuechterlein, K.H., Barch, D.M., Gold, J.M., Goldberg, T.E., Green, M.F., Heaton, R.K., 2004. Identification of separable cognitive factors in schizophrenia. Schizophr. Res. 72 (1), 29–39.

Nuechterlein, K.H., Green, M.F., Kern, R.S., Baade, L.E., Barch, D., Cohen, J., Essock, S., Fenton, W.S., Frese, F.J., Gold, J.M., Goldberg, T., Heaton, R., Keefe, R.S.E., Kraemer, H., Mesholam-Gately, R., Seidman, L.J., Stover, E., Weinberger, D., Young, A.S., Zalcman, S., Marder, S.R., 2008. The MATRICS consensus cognitive battery; part 1: test selection, reliability, and validity. Am. J. Psychiatry 165 (2), 203–213.

Patterson, T.L., Goldman, S., McKibbin, C.L., Hughs, T., Jeste, D.V., 2001. UCSD performance-based skills assessment: development of a new measure of everyday functioning for severely mentally ill adults. Schizophr. Bull. 27 (2), 235–245.

Ruse, S.A., Harvey, P.D., Davis, V.G., Atkins, A.S., Fox, K.H., Keefe, R.S., 2014. Virtual reality functional capacity assessment in schizophrenia: preliminary data regarding feasibility and correlations with cognitive and functional capacity performance. Schizophr. Res. Cogn. 1 (1), 21–26.

Silver, J.M., Koumaras, B., Chen, M., Mirski, D., Potkin, S.G., Reyes, P., Warden, D., Harvey, P.D., Arciniegas, D., Katz, D.I., Gunay, I., 2006. The effects of rivastigmine on cognitive function in patients with traumatic brain injury. Neurology 67 (5), 748–755.

Wechsler, D., 2003. Wechsler Intelligence Scale for Children. fourth edition. Pearson, San Antonio, TX.

Wechsler, D., 2008. Wechsler Adult Intelligence Scale. fourth edition. Pearson, San Antonio, TX.

Wykes, T., Huddy, V., Cellard, C., McGurk, S.R., Czobor, P., 2011. A meta-analysis of cognitive remediation for schizophrenia: methodology and effect sizes. Am. J. Psychiatry 168 (5), 472–485.