Published in final edited form as:

Schizophr Res. 2010 August ; 121(1-3): 24–31. doi:10.1016/j.schres.2010.04.016.

The Cognitive Assessment Interview (CAI): Development and Validation of an Empirically Derived, Brief Interview-Based Measure of Cognition

Joseph Ventura, Ph.D.¹, Steven P. Reise, Ph.D.², Richard S. E. Keefe, Ph.D.³, Lyle E. Baade, Ph.D.⁴, James M. Gold, Ph.D.⁵, Michael F. Green, Ph.D.^{1,6}, Robert S. Kern, Ph.D.^{1,6}, Raquelle Mesholam-Gately, Ph.D.⁷, Keith H. Nuechterlein, Ph.D.^{1,2}, Larry J. Seidman, Ph.D.⁷, and Robert M. Bilder, Ph.D.^{1,2}

¹Semel Institute for Neuroscience & Human Behavior, Geffen School of Medicine at UCLA ²UCLA Department of Psychology ³Department of Psychiatry and Behavioral Sciences, Duke University, Durham, North Carolina ⁴University of Kansas School of Medicine at Wichita, Kansas ⁵Maryland Psychiatric Research Center, University of Maryland School of Medicine, Baltimore, Maryland ⁶VA Greater Los Angeles Healthcare System, Los Angeles, California ⁷Massachusetts Mental Health Center, Public Psychiatry Division of the Beth Israel Deaconess Medical Center Department of Psychiatry, Harvard Medical School, Boston, Massachusetts

Abstract

Background—Practical, reliable “real world” measures of cognition are needed to supplement neurocognitive performance data to evaluate possible efficacy of new drugs targeting cognitive deficits associated with schizophrenia. Because interview-based measures of cognition offer one possible approach, data from the MATRICS initiative (n=176) were used to examine the psychometric properties of the Schizophrenia Cognition Rating Scale (SCoRS) and the Clinical Global Impression of Cognition in Schizophrenia (CGI-CogS).

Method—We used classical test theory methods and item response theory to derive the 10 item Cognitive Assessment Interview (CAI) from the SCoRS and CGI-Cogs (“parent instruments”). Sources of information for CAI ratings included the patient and an informant. Validity analyses

© 2010 Elsevier B.V. All rights reserved.

Address correspondence to: Joseph Ventura, Ph.D., UCLA Department of Psychiatry, 300 Medical Plaza, Room 2243, Los Angeles, CA. 90095, jventura@ucla.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Contributors

This project represents a collaborative effort with the research team from the National Institute of Mental Health supported Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) project. Data were obtained from the MATRICS Psychometric and Standardization Study (MATRICS PASS) led by the Co-Chairs of the Neurocognition Committee (Drs. Nuechterlein and Green). Drs. Ventura, Reise, and Bilder, conceived of the overall study design and data analysis plan. Dr. Ventura conducted literature searches, supervised the conduct of the study, and wrote the manuscript. Drs. Reise and Bilder conducted the data analysis, assisted in the presentation of the study results, and wrote sections of the manuscript. All authors have contributed to and approved the final manuscript.

Conflict of interest

The authors report no financial conflict of interest. However, this research was supported in part by an unrestricted grant from Pfizer, Inc., awarded to Joseph Ventura, Ph.D. (PI) and Robert M. Bilder, Ph.D. (Co-PI).

examined the relationship between the CAI and objective measures of cognitive functioning, intermediate measures of cognition, and functional outcome.

Results—The rater's score from the newly derived CAI (10-items) correlate highly ($r = .87$) with those from the combined set of the SCoRS and CGI-CogS (41 items). Both the patient ($r = .82$) and the informant ($r = .95$) data were highly correlated with the rater's score. The CAI was modestly correlated with objectively measured neurocognition ($r = -.32$), functional capacity ($r = -.44$), and functional outcome ($r = -.32$), which was comparable to the parent instruments.

Conclusions—The CAI allows for expert judgment in evaluating a patient's cognitive functioning and was modestly correlated with neurocognitive functioning, functional capacity, and functional outcome. The CAI is a brief, repeatable, and potentially valuable tool for rating cognition in schizophrenia patients who are participating in clinical trials.

Keywords

cognitive assessment; schizophrenia; item response theory; bifactor model; unidimensionality; intermediate outcomes

Introduction

Considering evidence which indicates that neurocognition is related to functional outcomes, neurocognitive deficits have become a target of treatment intervention (Carpenter and Gold 2002; Gold 2004; Green 2007). Findings from the National Institute of Mental Health supported Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) project evaluated which assessment methods can best be used to evaluate cognition-enhancing treatments for schizophrenia (Green, Nuechterlein et al. 2004; Nuechterlein, Barch et al. 2004; Green, Nuechterlein et al. 2008; Nuechterlein, Green et al. 2008). The MATRICS project was initiated because the U.S. Food and Drug Administration (FDA) has indicated that for any new pharmaceutical agents seeking approval as cognitive enhancers, improvements on a consensus, well validated objective cognitive battery would be necessary, but not sufficient. According to FDA officials, additional evidence from 'co-primary' measures also would be needed to corroborate that objectively measured change is, in fact, clinically meaningful in the "real world" (Green, Nuechterlein et al. 2008). As a result, additional outcome measures, such as interview-based measures would be required that are considered "face valid," functionally relevant, and/or clinically meaningful to the patients' lives.

In a separate set of guidelines for the pharmaceutical industry issued by the FDA, Patient-Reported Outcome (PRO) measures were deemed important for development of new cognitive enhancing medications (FDA, February, 2006)(Burke 2006). Further, these guidelines suggested that PRO measures should be included in clinical trials for new pharmaceutical agents because some treatment effects are known only to the patient. Further, data generated by a rater's use of a PRO instrument can provide evidence of a treatment benefit from the patient's perspective. However, for the data collected by a PRO to be meaningful, there should be empirical evidence that the PRO instrument effectively measures the particular construct being studied. Because ratings generated by patient self-report alone are correlated weakly or not at all with neurocognitive functioning in schizophrenia (van den Bosch and Rombouts 1998; Stip, Caron et al. 2003; Moritz, Ferahli et al. 2004; Prouteau, Verdoux et al. 2004; Hofer, Niedermayer et al. 2007), interview-based measures of cognition that are conducted by trained raters might be needed to provide reliable and valid information.

Although at a very early stage, research has been conducted on two interview-based measures of cognition, the Schizophrenia Cognition Rating Scale (SCoRS; (Keefe, Poe et al. 2006) and the Clinical Global Impression of Cognition in Schizophrenia (CGI-CogS; (Ventura, Cienfuegos et al. 2008). The three sources of information for rating items on the CGI-CogS and the SCoRS were: 1) the patient, 2) an informant (e.g., caregiver, social worker), and 3) a trained rater, who integrated the patient and informant data. The SCoRS (18-item scale) and the CGI-CogS (21 item scale) require a trained rater to evaluate the patient's and an informant's report of the patient's cognitive functioning. Analyses indicated that internal consistency and interrater reliability were good to excellent for both measures. In addition, the SCoRS rater composite was highly correlated with neurocognitive performance ($r = -.50$) and with functional outcome ($r = -.60$) (Keefe, Poe et al. 2006). The CGICogS rater composite was modestly correlated with neurocognitive performance ($r = -.32$) and highly correlated with functional outcome ($r = -.65$) (Ventura, Cienfuegos et al. 2008). The development of these forms of interview-based measures that are reliable and valid, easy to administer, and easily repeatable, potentially would promote their use to evaluate a patient's cognitive functioning in clinical trials.

One aim of the MATRICS Psychometric and Standardization Study (PASS) was to evaluate in a large sample the reliability, validity, and appropriateness for use in clinical trials of four potential co-primary measures: two of the four (the SCoRS and CGI-CogS) were interview-based measures of cognition. Data from MATRICS PASS indicated that both measures showed similar high test-retest reliability and small practice effects, supporting the potential utility of the SCoRS and CGI-CogS as repeated measures. Within the MATRICS PASS dataset, there were modest correlations with neurocognitive performance (SCoRS $r = -.31$, CGI-CogS $r = -.31$) and functional outcome (SCoRS $r = -.34$; CGI-CogS $r = -.30$) (Green, Nuechterlein et al. 2008). Considering all of the research thus far, the findings suggest the promise of interview-based assessments as co-primary measures of cognitive functioning.

In psychiatry, scale development and item selection are often guided by face validity and rarely by empirical methods. The SCoRS was rationally constructed based on content from experts in the field. The CGI-CogS was based on a model of cognitive deficits from the MATRICS project that specifies that there are seven separate, relevant domains of cognitive functioning for studies of cognitive enhancers (Nuechterlein, Barch et al. 2004). Yet, methodologically sophisticated statistical methods, such as item response theory (IRT), are available to aid in scale development (Embretson and Reise 2000). In fact, these methods were used on MATRICS PASS data to explore the dimensionality and the relative strengths and weaknesses of the SCoRS and CGI-CogS (Reise, Ventura et al. in press). These analyses indicated that there were advantages to each scale, but that a shortened version, if reliable and valid, would facilitate their repeated use in clinical trials.

The aim of this article is to describe how the Cognitive Assessment Interview (CAI) was developed using CTT and IRT from two existing interview-based measures of cognition, the CGICogS and the SCoRS. In addition, we examine the reliability and the concurrent validity of this brief, second-generation, co-primary measure.

Methods

The MATRICS Initiative involved extensive interactions among academic, NIMH, FDA, and pharmaceutical industry representatives (Green, Nuechterlein et al. 2008; Nuechterlein, Green et al. 2008). The data in this paper were obtained from the MATRICS Psychometric and Standardization Study (MATRICS PASS) led by the Co-Chairs of the Neurocognition Committee (Drs. Nuechterlein and Green). In MATRICS - PASS, data were collected on two occasions four weeks apart across five performance sites during this large study of

cognitive performance and functional outcome in schizophrenia patients. Human subject procedures were approved by each site's IRB. All subjects signed the approved informed consent form after the study was fully explained. The sample characteristics are described elsewhere (Green, Nuechterlein et al. 2008) and so will only be presented briefly (see Table 1).

MATRICES Assessments

The MATRICS PASS included performance-based measures of neurocognition (beta version of MCCB); (Nuechterlein, Green et al. 2008), interview-based measures of cognitive functioning, the Schizophrenia Cognition Rating Scale (SCoRS; (Keefe, Poe et al. 2006), the Clinical Global Impression of Cognition (CGI-CogS; (Ventura, Cienfuegos et al. 2008), a measure of functional capacity, the UCSD Performance-based Skills Assessment (UPSA; (Patterson, Goldman et al. 2001), ratings of clinical symptoms (BPRS; (Ventura, Green et al. 1993), and of community functioning (Birchwood Social Functioning Scale; (Birchwood, Smith et al. 1990). The measures addressed in the current study will be reviewed briefly here because they are more fully described elsewhere (Green, Nuechterlein et al. 2008).

A. Interview-based, Co-primary measures

1. Schizophrenia Cognition Rating Scale (SCoRS; (Keefe, Poe et al. 2006): The SCoRS is a 20-item (expanded for MATRICS-PASS) interview-based assessment of cognitive deficits and the degree to which they affect day-to-day functioning. A global rating is also generated. The items were developed to assess a variety of cognitive domains (e.g., memory, attention, problem solving) that were chosen because of the severity of impairment that has been shown in many patients with schizophrenia and the demonstrated relationship of these cognitive deficits to impairments in functional outcome. Two examples of items from the SCoRS are, "Do you have difficulty with remembering names of people you know?" and "Do you have difficulty following a TV show?" Each item is rated on a 4-point scale with higher ratings reflecting a greater degree of impairment.

2. Clinical Global Impression of Cognition in Schizophrenia (CGI-CogS; (Ventura, Cienfuegos et al. 2008): The CGI-CogS is similar in format to the SCoRS in that a patient and informant are interviewed for both measures and the interviewer provides a composite rating based on both sources of information. The CGI-CogS includes four major categories for evaluation: Activities of daily living, Neurocognitive state-Category severity, Global severity of cognitive impairment, and Global Assessment of Functioning. One key difference between the CGI-CogS and the SCoRS is that the CGI-CogS allows for ratings of cognition in specific cognitive domain, i.e., working memory, as well as overall cognition. The CGI-CogS uses a 7-point Likert scale with higher ratings indicating more impairment and interference with daily functioning.

Training and quality assurance for co-primary measures: For the interview-based assessments (SCoRS and CGI-CogS), initial training was provided on both scales in a one-day training session. Interviewers for these measures were selected from staff members at each site who had experience with semi-structured psychiatric interviews or symptom rating scales. The training was conducted by the developers of the scales.

B. University of California at San Diego Performance-Based Skills

Assessment (UPSA)—The UPSA is a functional capacity measure of five general skills that were previously identified as essential to functioning in the community: general organization, finance, social/communications, transportation, and household chores. The UCSD Performance-Based Skills Assessment involves role-play tasks that are administered as simulations of situations or events that the person may encounter in the community.

C. MATRICS Consensus Cognitive Battery (MCCB)—The MCCB, which has now been well described (Nuechterlein, Green et al. 2008), includes 10 tests from seven different cognitive domains: 1) Trail Making Test: Part A, 2) Brief Assessment of Cognition in Schizophrenia: Symbol-Coding, 3) Hopkins Verbal Learning Test – Revised, 4) Wechsler Memory Scale-III: Spatial Span 5) Letter-Number Span, 6) Neuropsychological Assessment Battery: Mazes, 7) Brief Visuospatial Memory Test – Revised, 8) Category Fluency (Animal Naming), 9) Mayer-Salovey-Caruso Emotional Intelligence Test: Managing Emotions, and 10) Continuous Performance Test – Identical Pairs.

D. Assessment of Symptoms and Community Functioning—Symptoms were assessed using the expanded version of the BPRS by raters trained to criterion levels of reliability (Ventura, Green et al. 1993). Variables from the Birchwood Social Functioning Scale (Birchwood, Smith et al. 1990) supplemented with work and school items from the Social Adjustment Scale (Weissman and Paykel 1974) were reduced through a principal components analysis into three domain scores (factor scores for work, social, and independent living) as well as a total score.

Development of the Cognitive Assessment Interview (CAI)

This study reports on the empirical development, which includes the reliability and validity, of a 10-item interview-based measure of cognitive functioning, titled the Cognitive Assessment Interview (CAI). The CAI is a subset of items derived from two parent instruments administered in MATRICS PASS, the SCoRS (20 items) and the CGI-CogS (21 items). Data from these parent instruments were subjected to a number of psychometric analyses (Reise, Ventura et al. in press). The coefficient alpha for CGI-CogS ratings was 0.95 with an average inter-item correlation of 0.47 (range from 0.15 to 0.65). The coefficient alpha for the SCoRS was 0.89 with an average inter-item correlation of 0.29 (range from 0.09 to 0.53). Using both individual and combined items, and traditional inspection of category frequencies, item means, and item-test correlations, rater data from these instruments were factor analyzed in three distinct ways: a) extracting a single factor (i.e., a unidimensional model), b) extracting multiple correlated dimensions, and c) through a bifactor framework where items were allowed to load on a general factor (cognitive deficit) as well as a number of secondary "group" factors. The conclusions drawn from those analyses were that both parent instruments had a strong common dimension underlying the items. In other words, a single latent factor was generally sufficient to explain item intercorrelations and there was evidence that both instruments are measuring the same common latent variable, i.e., cognitive deficit, and therefore could be shortened.

As a follow up to those analyses, an item response theory (IRT) model for polytomous items was fit to the data from the combined pool of CGI-CogS and SCoRS parent items. Although our psychometrics paper (Reise, Ventura et al. in press) focuses exclusively on the results for the rater data, in fact IRT models were additionally fit to the patient and informant data as well yielding similar results. The IRT results played a critical, but not determinative, role in selecting the "best" ten items that would form the CAI measure. Specifically, we inspected each "item information curve" (IIC) which indicated how informative, i.e., discriminating, each item was as a function of the latent trait, i.e., cognitive deficit. In that way, we were able to determine which scale items provided good precision at specific levels of the cognitive deficit. Second, again using the rater data, we conducted a simulated computerized adaptive test (CAT). We found that independent of an individual's level of cognitive deficit, the same set of 8 to 12 items (out of the 41-item pool) were indicated by CAT as informative. In addition, we found that trait level scores based on between 8 to 12 items were correlated very highly ($> .90$) with trait level scores based on administering the entire 41-item pool. These results are not too surprising given the highly unidimensional

nature of the CGI-CogS and SCoRS items demonstrated in the factor analyses we conducted.

The results of the CAT led us to conclude that the CAI needed to contain approximately 10 items to insure good measurement precision across a wide range of cognitive deficits. However, these empirical results were not the sole basis for selecting the final set of CAI items. We were interested in demonstrating not only that CAI had strong internal consistency, but also some reasonable degree of substantive breadth. The CAI items that were selected through this empirical process had strong psychometric properties (e.g., high factor loading in a unidimensional model, high item discrimination in the IRT analysis) as well as breadth. In particular, the empirically based item selection was guided by: a) whether the "good" psychometric properties found in the rater data were replicated in the patient and informant data, and b) the amount of information that an item provided in the middle of the cognitive deficit trait range (i.e., mild to moderate deficit). This was accomplished by rank ordering all of the SCoRS and CGI-CogS items (patient, informant, and rater) according to the item information curves across the three sources of information. Using that method, 3 items from the domain of Reasoning and Problem Solving (RPS) were in the top 8 – 12 items. We dropped one of the three RPS items so that there were comparable numbers of items represented in each domain, e.g., no more than two per MATRICS domain. None of the Spatial Learning and Memory items were represented in the top 8 – 12 items. This process resulted in a representation of six of the seven domains from MATRICS for a total of 10 CAI items (Table 2).

Results

Characteristics of the CAI

Ratings on the CAI are made based on information from the patient and an informant that is integrated in the rater score. Given that the CAI was not administered as a separate measure from the SCoRS and the CGI-CogS in MATRICS-PASS, all analyses are based on the 10 CAI patient, informant, and rater item means. The seven point rating scale is referenced to healthy people of similar educational and sociocultural background, with ratings of "1" reflecting healthy performance. Higher scores are associated with increasing cognitive deficits that impact on everyday functioning and/or increased need for support in performing those functions (Figure 1). The CAI includes scale items that assessed 6 of the 7 MATRICS cognitive domains (Table 2). The CAI correlates highly ($r = .87$) with the combined set of parent instruments (CGI-CogS and SCoRS; 41 items) (Table 2). The CAI includes the same global rating of cognitive function found in the CGI-CogS provided by the rater called the Global Assessment of Functioning - Cognition in Schizophrenia (GAF-CogS) that is rated on a 100-point scale.

Reliability of the CAI

Internal Consistency and Test-Retest Reliability—We found that the same 10 items that formed the CAI performed best across the different analyses, at several levels of the cognitive deficit construct, and had good internal consistency. Internal consistency for the CAI was very high as was the case for the parent instruments, the SCoRS and CGI-CogS. We calculated coefficient alpha for the CAI patient, informant, and rater ratings (all over .90; Table 3). Test-retest reliability for the CAI was assessed using data collected at baseline and at the one-month follow-up point (Table 3). Test-retest reliability for the CAI was high and the magnitude of the differences from baseline to the one month assessment was very small (patient effect size = -0.04 ; informant effect size = $-.04$; rater effect size = $-.08$). The effect size of change (**Cohen's d**) was for the patient $d = -0.04$, the informant $d = -0.04$, and for the rater score $d = -0.08$. In summary, test-retest, interrater, and internal consistency

were good to excellent in both “parent” scales (SCoRS and CGI-CogS). The CAI ratings show indications of being a reliable measure of cognitive deficits, but needed to be “psychometrically” validated. All of the results that follow are based on baseline relationships between the Cognitive Assessment Interview (CAI; mean of 10 items) and neurocognitive, functional capacity, and functional outcome data collected in MATRICS PASS.

Psychometric Validity of the CAI

The same methods that were used to validate the SCoRS and the CGI-CogS in MATRICS PASS were employed to examine the validity of the CAI. The correlations indicated that the CAI Patient and the CAI Rater score are highly related, as were the CAI Informant and CAI Rater score. Although the interviewer tended to be influenced by the informant, a CAI rater could conduct a valid assessment with the patient information alone (Table 4). Ratings based on Patient information and those based on Informant information were highly correlated with the Rater score (r 's = .82 and .95, respectively). Although the CAI Patient and CAI Informant ratings had a slightly lower correlation (r = .69), the CAI Rater score did not appear to gain a great deal of additional information from the availability of non-redundant Informant input.

An important consideration for evaluating the concurrent validity of co-primary measures is the degree to which they correlate with cognitive performance, functional capacity, and functional outcome, but not with psychiatric symptoms. To assess concurrent and divergent validity, the correlations for the CAI Patient, Informant, and Rater Score were examined at baseline with measures that included the: 1) MCCB neurocognitive assessment, 2) UCSD Performance Based-Skills Assessment (UPSA), 3) Birchwood self-reported level of functioning, and 4) BPRS Psychiatric symptoms. The key objective cognitive performance variable was defined as a single MCCB neurocognitive composite (NCS). This composite score was created by transforming individual test scores to z scores using the schizophrenia sample and then averaging across tests (Nuechterlein, Green et al. 2008). Correlations between the CAI rater score and the NCS were in the expected direction and significant (Table 5). The CAI was also moderately correlated with the UPSA (Table 5). Similar results were found when comparing the CAI ratings with the level of functional outcome (Table 5). Slightly lower values were noted for CAI ratings based solely on evaluation of the patient data. The CAI ratings were significantly correlated the BPRS positive symptoms factor (Table 6). However, the CAI was not significantly correlated with BPRS levels of negative symptoms or depression.

Step-wise multiple regression analyses were conducted to determine how much additional information is provided by the CAI patient rating as compared to the CAI informant rating in predicting relevant outcome variables that include neurocognition (MCCB), functional capacity (UPSA), and social functioning (Birchwood). In predicting the MCCB composite, when entered first, the CAI patient rating explains 6% of the variance and the CAI informant rating explains an additional 3% of the variance. In predicting the UPSA total score, when entered first the CAI patient rating explains 13% of the variance and the CAI informant rating explains an additional 4%. In predicting social functioning, when entered first the CAI patient rating explains 10% of the variance and the CAI informant does not explain any additional variance.

Predicting Functional Outcome with the CAI, MCCB, and the UPSA

A series of step-wise multiple regressions was performed to determine the relative contributions of different assessments of cognitive functioning, the MCCB, the UPSA, or the CAI, in the prediction of functional outcome as measured by the Birchwood Social

Functioning Scale (BSFS). With BSFS as the dependent variable, we forced entry of each of these instrument scores and then applied the stepwise procedure to enter the other variables. The findings are summarized in Figure 2, which shows how much variance in BSFS was accounted for each of the cognitive assessments, depending on which was entered first. The figure shows that when CAI was entered first, the CAI rating explains 10%, followed by MCCB composite 3%, and finally UPSA total score at 0% of variance in BSFS. When the MCCB composite was entered first it explains 8%, then CAI rating 5% and finally UPSA total at 0%. When UPSA total is entered first, it explains 6% of variance in BSFS, then the CAI rating adds 5% and finally MCCB composite 2%. Taken together, these results suggest that: (1) the CAI is more strongly related to functional outcomes than either the MCCB or UPSA; (2) the UPSA does not appear to add any unique variance above that contributed by MCCB and CAI; and (3) the CAI does contribute unique variance in predicting outcome beyond that attributable to both MCCB and UPSA.

Discussion

Interview-based measures of cognition, such as the CAI are now undergoing a second cycle of evolution that is attempting to find reliable and valid ways to rate cognitive functioning. The CAI is a semi-structured, interview-based measure of cognitive functioning that requires administration by a trained rater. The CAI includes a patient's report of his or her cognitive functioning and that of an informant, both of which are evaluated to create a rater composite. We used classical test theory (CTT) and item response theory (IRT) methods to examine the psychometric properties of two parent instruments, the SCoRS and CGI-CogS. Given the unidimensional nature of the data we concluded that these instruments could be shortened without loss of measurement precision. The most informative items were selected through empirical means from those first generation parent measures to comprise the 10 CAI items. The CAI was designed to yield ratings of severity of cognitive deficits and their impact on community functioning. We found that the CAI shows indications of being a reliable and valid measure of cognitive functioning, functional capacity, and functional outcome comparable to either of the "parent" instruments. The CAI shows promise as a co-primary measure for clinical trials designed to study new cognitive enhancing agents. However, the CAI is not meant to be a substitute for objective cognitive testing.

Second-generation interview-based measures of cognition such as the CAI potentially have several theoretical and practical advantages that could ultimately facilitate their utility in clinical trials. Interview-based measures might be tapping into an important dimension of cognitive deficits that overlaps with, but is also independent of objective tests of cognition. The CAI and the UPSA (a performance-based measure) have both been proposed as co-primary endpoints in clinical trials. The criterion for final selection was established as the correlation with objective cognitive testing. Yet, correlations between the MCCB and the UPSA tend to be higher than correlations between the MCCB and interview-based assessments such as the CAI. This is likely to be reflecting an element of shared method variance given that both the MCCB and the UPSA depend on performance of tasks with similar cognitive demands. In this way, the CAI might offer a rating of cognitive functioning that is not as redundant with objectively measured cognition as is the UPSA. On the question of how well these scales predict outcome, the MATRICS PASS data shows that the CAI correlates as strongly with functional outcome ($r = -.32$) as does the UPSA ($r = .23$) (Green, Nuechterlein et al. 2008). However, the CAI might have higher correlations with functional outcome scales partly because both involve questions about daily functioning. In any event, our multiple regression results suggest that the CAI is at least as closely related to real world functional outcomes as are objective measures of cognition (MCCB) and functional capacity (UPSA).

Whether interview-based data from patients alone can be used to assess cognition is a matter of some debate. According to Green and colleagues (Green, Nuechterlein et al. 2008) data collection was nearly complete for the patient interview-based assessments. However, about 14% of the informant interviews (across both assessment periods) were missing. The potential difficulty in locating informants might make it challenging to rely on measures that require informant data in schizophrenia clinical trials in contrast to dementia trials for which informants are typically available. Judging from the CGI-CogS and SCoRS correlations in the PASS data set, informant data appears to have added only incrementally to the reliability and the validity of the rater based scores. Fortunately, the CAI ratings based on patient information alone showed high test-retest reliability, high internal consistency, and similar validity statistics. Our interpretation of the regression analyses is that the CAI Patient rating alone provided a good deal of information for predicting cognitive functioning, functional capacity, and social functioning. In contrast to prior work suggesting that patient selfreport measures lack validity with respect to relationships with objective neurocognitive test scores or real-world outcomes, the findings reported here suggest that the CAI does provide reliable and meaningful information about the patient's functioning (van den Bosch and Rombouts 1998; Stip, Caron et al. 2003; Moritz, Ferahli et al. 2004; Prouteau, Verdoux et al. 2004).

As with all studies, there are limitations. The reliability and validity shown here must be interpreted in the context of the MATRICS PASS sample. Larger or more representative samples could yield different results regarding the use of interview-based measures. Also, there will always be some doubt about whether patients with schizophrenia can reliably describe their own cognitive impairment. Future work should examine whether patient reported outcomes might show reduced reliability and validity in patients with very poor cognitive functioning.

A critical future direction for the Cognitive Assessment Interview includes determining empirically whether this instrument is sensitive to clinically meaningful change in response to an intervention. To address this issue will require either pharmaceutical or cognitive remediation interventions that can reliably improve cognitive functioning in schizophrenia patients. For the benefit of our patients, we hope such new treatments will be forthcoming.

Acknowledgments

The findings from these analyses were presented in part at the 12th bi-annual meeting of the International Congress on Schizophrenia Research, San Diego, California, March 28 – April 1, 2009; Ventura, J., Reise, S.R., Bilder, R.M., and Keefe, R.S., Development of an Interview-Based “Co-Primary” Measure of Cognition

This research was supported by an investigator initiated grant from Pfizer, Inc., and from an NIMH grant 1R21MH073971 that were awarded to Joseph Ventura, Ph.D. Funding for the MATRICS Initiative was provided through contract N01MH22006 from NIMH to the University of California, Los Angeles, Dr. Marder, Principal Investigator; Dr. Green, Co-principal investigator; Dr. Fenton, Project Officer. Funding for MATRICS PASS came from an option (Dr. Green, principal investigator; Dr. Nuechterlein, co-principal investigator) to the NIMH MATRICS Initiative. This research was also supported in part by National Institute of Mental Health Grants MH37705 (P.I.: Keith H. Nuechterlein, Ph.D.), and P50 MH066286 (P.I.: Keith H. Nuechterlein, Ph.D.).

The authors would like to thank Sarah Wilson, M.A., for her assistance in organizing and preparing the data for analysis.

Role of funding source

The funding sources did not play role in the study design, implementation of data analysis, results, or publication of this paper.

References

- Birchwood M, Smith J, et al. The social functioning scale: the development and validation of a new scale of social adjustment for use in family intervention programs with schizophrenic patients. *British Journal of Psychiatry*. 1990; 157:853–859. [PubMed: 2289094]
- Burke, L.; Stifano, Toni; Dawisha, Sahar. Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. Rockville, MD: U.S. Department of Health and Human Services, FDA; 2006. p. 1-32.
- Carpenter WT, Gold JM. Another view of therapy for cognition in schizophrenia. *Biological Psychiatry*. 2002; 51(12):969–971. [PubMed: 12062880]
- Embretson, S.; Reise, S. Item response theory for psychologists. Lawrence Erlbaum Associates; 2000.
- Gold JM. Cognitive deficits as treatment targets in schizophrenia. *Schizophrenia Research*. 2004; 72(1):21–28. [PubMed: 15531404]
- Green M. Cognition, drug treatment, and functional outcome in schizophrenia: a tale of two transitions. *American Journal of Psychiatry*. 2007; 164(7):992. [PubMed: 17606645]
- Green MF, Nuechterlein KH, et al. Approaching a consensus cognitive battery for clinical trials in schizophrenia: The NIMH-MATRICES conference to select cognitive domains and test criteria. *Biological Psychiatry*. 2004; 56:301–307. [PubMed: 15336511]
- Green MF, Nuechterlein KH, et al. Functional Co-Primary Measures for Clinical Trials in Schizophrenia: Results From the MATRICES Psychometric and Standardization Study. *American Journal of Psychiatry*. 2008
- Green MF, Nuechterlein KH, et al. Functional Co-Primary Measures for Clinical Trials in Schizophrenia: Results From the MATRICES Psychometric and Standardization Study. *Am J Psychiatry*. 2008 appi.ajp.2007.07010089.
- Hofer A, Niedermayer B, et al. Cognitive impairment in schizophrenia: clinical ratings are not a suitable alternative to neuropsychological testing. *Schizophrenia research*. 2007; 92(1–3):126–131. [PubMed: 17350807]
- Keefe RS, Poe M, et al. The Schizophrenia Cognition Rating Scale: an interview-based assessment and its relationship to cognition, real-world functioning, and functional capacity. *Am J Psychiatry*. 2006; 163(3):426–432. [PubMed: 16513863]
- Moritz S, Ferahli S, et al. Memory and attention performance in psychiatric patients: lack of correspondence between clinician-rated and patient-rated functioning with neuropsychological test results. *Journal of the International Neuropsychological Society*. 2004; 10(04):623–633. [PubMed: 15327740]
- Nuechterlein K, Green M, et al. The MATRICES Consensus Cognitive Battery, part 1: test selection, reliability, and validity. *American Journal of Psychiatry*. 2008; 165(2):203. [PubMed: 18172019]
- Nuechterlein KH, Barch DM, et al. Identification of separable cognitive factors in schizophrenia. *Schizophrenia Research*. 2004; 72(1):29–39. [PubMed: 15531405]
- Nuechterlein KH, Barch DM, et al. Identification of separable cognitive factors in schizophrenia. *Schizophrenia Research*. 2004; 72:29–39. [PubMed: 15531405]
- Patterson TL, Goldman S, et al. UCSD Performance-Based Skills Assessment: Development of a new measure of everyday functioning for severely mentally ill adults. *Schizophrenia Bulletin*. 2001; 27(2):235–245. [PubMed: 11354591]
- Prouteau A, Verdoux H, et al. Self-assessed cognitive dysfunction and objective performance in outpatients with schizophrenia participating in a rehabilitation program. *Schizophrenia research*. 2004; 69(1):85–91. [PubMed: 15145474]
- Reise SP, Ventura J, et al. Bifactor and Item Response Theory Analyses of Interviewer Report Scales of Cognitive Functioning in Schizophrenia. *Psychological Assessment*. (in press).
- Stip E, Caron J, et al. Exploring cognitive complaints in schizophrenia: the subjective scale to investigate cognition in schizophrenia* 1. *Comprehensive psychiatry*. 2003; 44(4):331–340. [PubMed: 12923712]
- van den Bosch R, Rombouts R. Causal Mechanisms of Subjective Cognitive Dysfunction in Schizophrenic and Depressed Patients. *The Journal of nervous and mental disease*. 1998; 186(6): 364. [PubMed: 9653421]

- Ventura J, Cienfuegos A, et al. Clinical global impression of cognition in schizophrenia (CGI-CogS): Reliability and validity of a co-primary measure of cognition. *Schizophrenia Research*. 2008; 106(1):59–69. [PubMed: 17900866]
- Ventura J, Green MF, et al. Training and quality assurance on the Brief Psychiatric Rating Scale: the "drift busters". *International Journal of Methods in Psychiatric Research*. 1993; 3:221–224.
- Weissman, M.; Paykel, E. *The Depressed Woman: A Study of Social Relationships*. Chicago, IL: University of Chicago Press; 1974.

DOMAIN: Attention/Vigilance																								
3. Problems sustaining concentration over time (without distraction)?																								
Do you have trouble concentrating? Do you take breaks frequently? Do you have trouble paying attention while reading, listening to the radio or watching television, long enough to read/listen/see a whole article/chapter/program?																								
Patient Examples:										Informant Examples:														
Patient								Informant								Rater								
N/A	1	2	3	4	5	6	7	N/A	1	2	3	4	5	6	7	N/A	1	2	3	4	5	6	7	
Severity Anchor Points																								
N/A = Rating not applicable, or insufficient information					1. Normal, not at all impaired					2. Minimal cognitive deficits but functioning is generally effective					3. Mild cognitive deficits with some consistent effect on functioning									
4. Moderate cognitive deficits with clear effects on functioning					5. Serious cognitive deficits which interfere with day-to-day functioning					6. Severe cognitive deficits that jeopardize independent living					7. Cognitive deficits are so severe as to present danger to self/others									

Figure 1.
Example of a CAI Item, with Cognitive Domain Name, Domain Definition, Probe Questions, Rating Scale, and Anchor Point Definitions

Figure 2. Predicting Functional Outcome from Cognitive Assessments

Results of stepwise regressions with Birchwood Social Functioning Scale (BSFS) as the dependent variable, and either the CAI, MCCB or UPSA scores entered first (each clustered group of bars shows results for one regression). The individual bars within each cluster show how much variance in BSFS is shared with each of the instruments in the resulting equation. Note that together, the three cognitive assessment scales explain about 13% of variance in BSFS scores; CAI showed the largest contribution and also added unique variance to MCCB and UPSA scores; UPSA did not add unique variance beyond that contributed by CAI and MCCB.

Table 1Sample Characteristics n = 176¹)

Age (years)	44.0(11.2)	
Education	12.4 (2.4)	
Sex	76% male	
Marital Status	Single Married Divorced Widowed Separated	61.4% 9.7 % 23.9 % 2.3 % 2.8%
Ethnicity	Caucasian African Am Latino Other	59% 29% 6% 5%
Diagnosis	Schizophrenia Schizoaffective, Depressed type	86% 14%
BPRS symptoms ² Positive Negative	2.55(1.26) 1.99(0.86)	

¹ For study analyses that involved Informant data, the number of subjects was typically lower than 176.

² Both types of symptoms were in the mild range and typically below clinically significant levels.

Table 2

The 10 Items of the Cognitive Assessment Interview (CAI) and their Correlation with the Combined 41-Item Total Score from the SCoRS and CGI-CogS (N = 172)

	Item-scale correlation
Working Memory	
1. Difficulty maintaining newly learned information	.717
2. Difficulty performing on the spot mental manipulations	.710
Attention Vigilance	
3. Difficulty sustaining concentration over time	.668
4. Difficulty focus on selected information (without distraction)	.740
Verbal Learning and Memory	
5. Trouble Learning and remembering verbal information	.696
6. Difficulty recalling recent events	.645
Reasoning and Problem Solving	
7. Lack of flexibility in generating alternative plans when needed	.705
8. Problems in situations requiring judgment	.773
Speed of Processing	
9. Performs tasks slowly	.686
Social Cognition	
10. Difficulty appreciating another person's intentions or point of view	.722
Total of all CAI items	.872

Table 3

Reliability Data for the 10-item CAI (N=172)

10-item CAI	Cronbach's Alpha (baseline)	Test-Retest Reliability (ICC)
CAI Patient	0.91	0.84
CAI Informant	0.92	0.91
CAI Rater	0.96	0.89

**
= p < .01,

*
= p < .05

Table 4

Relationships among the 10-item CAI for the Patient, Informant, and Rater Score (Pearson r) (n = 170)

10-item CAI	CAI Patient	CAI Informant	CAI Rater
CAI Patient	1.0	0.69**	0.82**
CAI Informant		1.0	0.95**
CAI Rater			1.0

**
= p < .01,

*
= p < .05

Table 5

Correlations Between the SCoRS, CGI-CogS, CAI with Cognition, Functional Capacity, and Functional Outcome (N=172)

Co-primary Measure	Overall Cognition (MCCB)	Functional Capacity (UPSA)	Functional Outcome (Birchwood)
SCoRS			
Patient	-0.18	-0.36 **	-0.23
Informant	-0.21	-0.38 **	-0.24
Rater Composite	-0.31	-0.42 **	-0.34
CGI-CogS			
Patient	-0.30	-0.44 **	-0.31
Informant	-0.24	-0.40 **	-0.27
Rater Composite	-0.31	-0.48 **	-0.30
CAI-10 items			
CAI Patient	-0.28 **	-0.42 **	-0.31 **
CAI Informant	-0.30 **	-0.40 **	-0.27 **
CAI Rater	-0.32 **	-0.44 **	-0.32 **

Note: Analyses on the relationship between SCoRS and CGI-CogS with overall cognition and functional outcome are from Green et al. (2008).

** = $p < .01$,

* = $p < .05$

Table 6

Relationships between the 10-item CAI and Symptoms (Pearson r) at Baseline (N=170)

10-item CAI	Symptoms		
	Positive	Negative	Depression
CAI - Patient	.28**	.09	.20*
CAI - Informant	.27**	.15	.04
CAI - Rater	.28**	.11	.08

**
= p < .01,

*
= p < .05