



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

Schizophrenia Research xx (2007) xxx–xxx

SCHIZOPHRENIA  
RESEARCH[www.elsevier.com/locate/schres](http://www.elsevier.com/locate/schres)

## Clinical global impression of cognition in schizophrenia (CGI-CogS): Reliability and validity of a co-primary measure of cognition

Joseph Ventura<sup>a,\*</sup>, Angel Cienfuegos<sup>a,b</sup>, Oren Boxer<sup>a</sup>, Robert Bilder<sup>a,c</sup>

<sup>a</sup> UCLA Department of Psychiatry and Biobehavioral Sciences, 300 Medical Plaza, Room 2243, Los Angeles, CA 90095, USA

<sup>b</sup> Department of Veterans Affairs VISN 22 Mental Illness Research, Education, and Clinical Center; Greater Los Angeles VA Health Care System, 11301 Wilshire Blvd., 210A Los Angeles, CA 90073, USA

<sup>c</sup> UCLA Department of Psychology, Franz Hall, Los Angeles, CA 90095, USA

Received 2 January 2007; received in revised form 19 July 2007; accepted 23 July 2007

### Abstract

**Background:** Cognitive deficits are core features of schizophrenia that have been associated reliably with functional outcomes and now are a focus of treatment research. New rating scales are needed to complement current psychometric testing procedures, both to enable wider clinical use, and to serve as endpoints in clinical trials.

**Methods:** Subjects were 35 schizophrenia patient-and-caregiver pairs recruited from the UCLA and West Los Angeles VA Outpatient Psychiatry Departments. Participants were assessed with the Clinical Global Impression of Cognition in Schizophrenia (CGI-CogS), an interview-based rating scale of cognitive functioning, on 3 occasions (baseline, 1 month, and 3 months). A computerized neurocognitive battery (Cogtest), an assessment of functioning, and symptom measures were administered at two occasions (baseline and one month).

**Results:** The CGI-CogS ratings generally showed a high level of internal consistency (Cronbach's alpha = .69 to .96), adequate levels of inter-rater reliability (ICC's = .71 to .80), and high test–retest stability (ICC's = .92 to .95). Correlations of caregiver and rater global (but not “patient only rating”) CGI-CogS ratings with neurocognitive performance were in the moderate range ( $r$ 's = -.27 to -.48), while most of the correlations with functional outcome were moderate to high ( $r$ 's = -.41 to -.72). In fact, the CGI-CogS ratings were significantly more correlated with Social Functioning than were objective neurocognitive test scores ( $p$  = .02) and showed a trend in the same direction for predicting Instrumental Functioning ( $p$  = .06). We found moderate correlations between CGI-CogS global ratings and PANSS positive ( $r$ 's = .36 to .49) and SANS negative symptoms ( $r$  = .41 to .61), but not with BPRS depression ( $r$ 's = .11 to .13).

**Conclusions:** An interview-based measure of cognition demonstrated high internal consistency, good inter-rater reliability, and high test–retest reliability. Caregiver ratings appear to add important clinical information over patient-only ratings. The CGI-CogS showed moderate validity with respect to neurocognitive performance and functional outcome, and correlations of CGI-CogS with functional outcomes were stronger than correlations of objective neurocognitive performance with functional outcomes. The CGI-CogS appears to offer a reliable and valid method for clinical rating of cognitive deficits and their impact on everyday functioning in schizophrenia.

© 2007 Elsevier B.V. All rights reserved.

**Keywords:** Schizophrenia; Cognitive assessment; Interview-based measure; Caregiver; Functional outcome; Co-primary measure

\* Corresponding author. Tel.: +1 310 2065225; fax: +1 310 2063651.

E-mail address: [jventura@ucla.edu](mailto:jventura@ucla.edu) (J. Ventura).

## 1. Introduction

Cognitive deficits are a defining feature of schizophrenia that are closely related to functional outcomes and activities of daily living (Green, 1996; Green et al., 2000, 2004; Green and Nuechterlein, 1999). There has been increasing attention on developing effective treatment interventions both psychopharmacological and non-pharmacological, to remediate these core cognitive deficits (Carpenter, 2004; Gold, 2004; Silverstein and Wilkness, 2004). The value of cognitive assessment in clinical trials has been supported by the NIMH MATRICS project, which identified seven domains of cognition important for schizophrenia; vigilance/attention, working memory, processing speed, verbal learning and memory, visual learning and memory, reasoning and problem solving, and social cognition (Green and Nuechterlein, 2004; Nuechterlein et al., 2006a,b). However, there are both theoretical and practical limitations to the primary reliance on measuring improvement in cognition through objective, cognitive performance tests. In fact, there is a significant gap between identifying schizophrenia patient deficits in objective testing situations and in understanding their impact on patient's lives and everyday functioning. The FDA has indicated that neuropsychological tests alone are inadequate for the evaluation of improvement of a patient's cognitive functioning. Guidelines for the development of interview-based measures of cognition can be found in a recent document from the FDA that emphasize sensitivity to patient defined change and the importance of developing patient oriented outcome measures (February, 2006).

Clinicians who work with schizophrenia patients and their families need cognitive assessment tools that are easy to administer and can validly evaluate change in cognitive functioning. In addition, researchers need a tool that can assess cognitive skills and abilities that are directly linked with behaviors associated with a patient's daily functioning, e.g., paying attention at work. Patients, families, and clinicians are seeking improved real world outcomes and a better quality of life rather than improved scores on a battery of neuropsychological tests. But, there are limitations to neuropsychological tests used by researchers in clinical trials and by clinicians in neuropsychological assessments. For example, patients who score well on objective memory tests might still have problems with memory that interfere with their daily functioning. In addition, practical limitations such as the amount of prior training, administration and scoring time, practice effects, and validity issues associated with interpretation ultimately reduce the feasibility of objective tests. To measure change in cognition during clinical trials and clinical practice, new

assessment tools are needed. Interview-based assessments might offer some promise in overcoming the practical limitations that exist regarding the primary use of objective tests of neurocognitive functioning.

The use of interview-based measures of cognition that includes a clinician's judgment potentially improves validity because purely subjective impressions of cognitive function alone are only weakly or not at all correlated with objective neurocognitive testing (Moritz et al., 2004; Prouteau et al., 2004; Harvey et al., 2001; van den Bosch and Rombouts, 1998). Nonetheless, purely objective impressions by a clinician alone cannot be used without being evaluated through a patient's or a caregiver's report. Designing an interview-based cognitive functioning scale almost certainly could follow the dementia model and incorporate not only patient report but also caregiver collateral information (Keefe et al., 2006). Of course, the inclusion of caregiver input does not assure validity in the clinical assessment of any disorder. Reliable and valid interview-based assessments of cognition might be dependent upon obtaining the patient's report, a caregiver's report, and a clinical evaluation of both sources of information. In fact, some studies have suggested that a patient's complaints about cognitive functioning should be used to assist in the evaluation of a patient's current level of cognitive functioning (Prouteau et al., 2004).

The current study evaluated the inter-rater reliability and validity of an interview-based co-primary measure of cognition called the Clinical Global Impression of Cognition in Schizophrenia (CGI-CogS) that was administered to schizophrenia patients and their caregivers. We examined the inter-rater reliability and validity of the CGI-CogS according to the criteria established by the FDA panel and the MATRICS committee: a) good test–retest reliability, b) demonstrated associations with cognitive performance measures, and c) demonstrated associations with community functioning. To examine reliability of the CGI-CogS, we evaluated the internal consistency using Cronbach's alpha, tested the inter-rater reliability among independent raters, and assessed the test–retest reliability. To examine validity of the CGI-CogS, we evaluated correlations with an objective test of cognitive functioning, psychiatric symptoms, and functional outcome.

## 2. Methods

### 2.1. Subjects

The sample consisted of 35 patient and caregiver pairs that were recruited from ongoing research projects at the UCLA Aftercare Research Program and the West

Table 1  
Basic demographic characteristics of patients and caregivers

	Patients (N=35)	Caregivers (N=18)
Age (years)	38.5 (12.0)	51.5 (14.7)
Education (years)	13.6 (2.3)	16.9 (2.6)
Sex (M, F)	23, 12	2, 33
Handedness (#R, #L)	32, 3	17, 1
Marital status	30 (86%) Single	11 (61%) Single
	3 (9%) Married	4 (22%) Married
	2 (6%) Divorced/ Widowed	3 (17%) Divorced/ Widowed
Ethnicity	10 (29%) Caucasian	8 (44%) Caucasian
	12 (34%) African Am	4 (22%) African Am
	6 (17%) Latino	3 (17%) Latino
	5 (14%) Asian	3 (17%) Asian
	2 (6%) Mixed	0 (0%) Mixed
Interview time (minutes)	33.1 (5.8)	32.3 (6.6)

Los Angeles VA Medical Center (see Table 1). The patients were diagnosed with schizophrenia (82%), schizoaffective disorder (3%), or schizophreniform disorder (15%). These outpatients were 66% male, had a mean age at study entry of 38.5 ( $\pm$ 12) years, and a mean educational achievement of 13.6 ( $\pm$ 2.3) years (see Table 1). The caregivers ( $n$ =18) were 94% female, had a mean age of 51.5 ( $\pm$ 14.7) years, and a mean educational achievement of 16.9 ( $\pm$ 2.6) years. Overall, the patients were of an age, educational achievement, and sex distribution typical of those reported in the schizophrenia research literature.

Although we enrolled 35 patients and caregiver pairs, please note that for a several patients a UCLA or VA case manager served as the caregiver. For most analyses, there are 33 complete data points at baseline and one month. We also completed 28 CGI-CogS assessments on patients and caregivers at the 3-month follow-up point. However, data from the 3-month follow-up are not provided in this paper. Patients were only included in this study if they satisfied objective criteria for clinical stability for the 8 weeks preceding study entry. Clinical stability criteria were defined as *not* being in a state of acute exacerbation or relapse during the 8 weeks before study entry (Nuechterlein et al., 2006a,b). All participants gave written informed consent after verbal and written information about this project was provided.

## 2.2. Procedures

### 2.2.1. Background and development of the CGI-CogS

The Clinical Global Impression of Cognition in Schizophrenia (CGI-CogS) was modeled after the Clinical Interview-Based Impression of Severity (CIBIS) and Clinical Interview-Based Impression of Change with

input from caregivers (CIBIS/CIBIC+), widely used in dementia research. The CGI-CogS eliminated ratings of psychotic symptoms, such as delusions and hallucinations, which are included in the CIBIS/CIBIC+. While in dementia the rating of psychotic symptoms might be a valid index of severity, the rating of psychosis would be redundant in schizophrenia treatment research with traditional rating scales. The CGI-CogS goes beyond CIBIS/CIBIC+ in rating separately the seven neurocognitive domains identified in the MATRICS project (see [www.matrics.ucla.edu](http://www.matrics.ucla.edu)) that include: Working Memory, Attention/Vigilance, Verbal Learning and Memory, Visual Learning and Memory, Reasoning and Problem Solving, Speed of Processing, and Social Cognition.

### 2.2.2. A co-primary interview-based measure of cognition

2.2.2.1. Clinical Global Impression of Cognition in Schizophrenia (CGI-CogS; Bilder et al., 2003). The CGI-CogS has a general background section and includes two major categories for evaluation: Activities

CGI-CogS Global, Category, Domain, and Item Structure

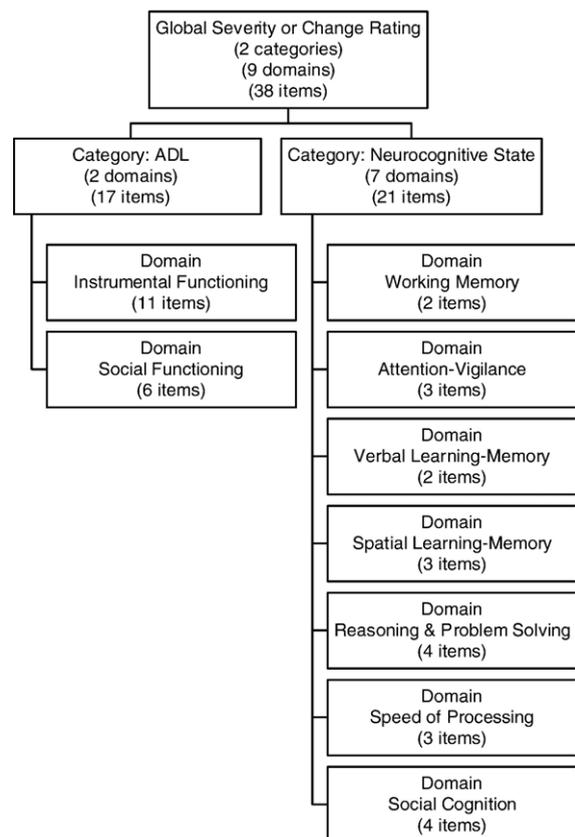


Fig. 1. Structure of the CGI-CogS domains and Items.

DOMAIN: Working Memory																												
1. Difficulty maintaining newly learned verbal information in mind for brief periods (long enough to use)?																												
<b><i>Do you forget names of people you just met? Do you have trouble recalling telephone numbers you hear? Do you have trouble remembering what your Dr. just said during visits? Do you find you need to write down information to remember?</i></b>																												
Patient Examples:											Caregiver Examples:																	
Patient							Caregiver							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							

Fig. 2. Example of CGI-CogS Item, Domain Name, Domain Definition, Probe Questions, Rating Scale, and Anchor Point Definitions.

of Daily Living and Neurocognitive State (see Fig. 1). These categories are further broken down into domains: Activities of Daily Living has two domains and Neurocognitive State has seven domains (see Fig. 1). The seven domains in the Neurocognitive State category explicitly follow the domains identified in the MATRICS project. Within each domain there are a series of items ( $n$  of items in each domain is listed in Fig. 1); each item contains one or more “probes” that are used by raters as needed to determine the score on that item. Each item is evaluated across three sources: (1) patient; (2) caregiver, and (3) all available sources of information combined as evaluated by the rater. Each item uses a seven point rating scale with operational anchor points, and a place to record verbatim responses (see example item; Fig. 2). The seven point rating scale is referenced to healthy people of similar educational and sociocultural background, with ratings of “1” reflecting healthy performance, and higher scores associated with increasing difficulty, increasing impact on everyday functioning, and/or increased need for support in performing these functions. In addition to individual item ratings, several “global” ratings were made, the most important being the ADL and Neurocognitive State Category scores, and the Global Assessment of Functioning score. Further details are provided below. The complete instrument including instruction manual and rating booklet is available for downloading at the UCLA Center for Neurocognition and Emotion in Schizophrenia web site (<http://www.schizophrenia.ucla.edu/assessment/>).

**2.2.2.2. General background section.** In the CGI-CogS “General” category, the interviewer records information about personal and medical background including current state that may help explain ratings of

cognition in the individual item or domains, e.g., recent head injury.

**2.2.2.3. Activities of Daily Living.** The section on Activities of Daily Living (ADL) contains two subscales, Instrumental Functioning (IF; 11 items) and Social Functioning (SF; 6 items) for an ADL total of 17 items. The section on Instrumental Functioning included items such as, hygiene, hobbies, and household chores. The section on Social Functioning included items such as, socializing with family and peers, community activities, and dating. The primary variables of interest were ratings on the *Instrumental Functioning* and *Social Functioning* subscales that for some analyses were combined into a single category *Activities of Daily Living*.

**2.2.2.4. Neurocognitive deficit severity.** This section includes all 21 individual CGI-CogS cognitive items with a set of individualized questions and probes that were meant to capture cognitive deficits. The CGI-CogS involves rating of the patient’s cognitive functioning at the Item, Domain, and Category levels on a 7-point scale where “1” represents completely “normal” function and “7” represents “severe” impairment. The Domain ratings were made after scoring all Items within a Domain, e.g., working memory. The global ratings are made after scoring all seven Domains for each Category. The primary variables of interest were two global ratings, *Neurocognitive State-Category Severity (NS-CS)* and *Global Severity of Cognitive Impairment (GS-CI)*.

**2.2.2.5. Global assessment of cognitive functioning.** The CGI-CogS interview allows the rater to assess global cognitive functioning using the Global Assessment of Functioning — Cognition in Schizophrenia

(GAF-CogS) scale, which is rated on a 100-point scale. The GAF-CogS is intended to supplement the CGI-CogS global severity ratings, and parallels the DSM-IV GAF scale. The anchors for the GAF-CogS instruct the rater to evaluate the extent of functional impairment associated with cognitive impairment rather than more traditionally rated psychiatric symptoms. A high level of functioning with no cognitive impairment is rated “100” as compared to a rating of “1” which indicates the lowest level of functioning that is associated with severe levels of cognitive impairment. The primary variable of interest, as scored for the patient, caregiver, and rater was the *Global Assessment of Functioning-Cognition in Schizophrenia (GAF-CogS)* score.

### 2.2.3. Objective tests of neurocognitive functioning

**2.2.3.1. Cogtest.** Cogtest is a computerized neurocognitive assessment battery that includes various individual neurocognitive tests. The tests use a touch screen for some participant responses, while the mouse or keyboard is used for other responses. The tests included in this study are: CPT AX-version, Spatial Working Memory Test, Auditory Digit Span, Auditory Number Sequencing, Go/No-Go Test, Set Shifting Test, Strategic Target Detection Test, Symbol Digit Substitution Test, Word List Memory Test, and Face Memory Test. Further descriptions of these tests and citations are available (<http://www.cogtest.com>). The analysis of concurrent validity in the current study for the Cogtest scores first involved inspection of all Cogtest variables. Then, a computation of standardized scores on the Cogtest variables was performed to create a Neurocognitive Composite Score (NCS).

### 2.2.4. Psychiatric Symptom Rating Scales

**2.2.4.1. The 24-item Brief Psychiatric Rating Scale (Ventura et al., 1993).** This version of the classic BPRS was used to generate scores on four factors (Positive Symptoms, Negative Symptoms, and Anxiety/Depression; Ventura et al., 2000).

**2.2.4.2. Scale for the Assessment of Negative Symptoms (SANS; Andreasen, 1984).** The SANS contains 20 items within five domains: Affective flattening; Alogia, Avolition–Apathy; Anhedonia–Asociality; and Attention.

**2.2.4.3. The Positive and Negative Syndrome Scale (PANSS; Kay et al., 1987).** The PANSS contains 30 items and yields three domain scores; positive psychosis, negative symptoms, and general psychopathology.

### 2.2.5. Data collection protocol

Data were collected at the UCLA Aftercare Research Program and the Veteran’s Administration Department of Outpatient Psychiatry. All symptom ratings were made using established training programs and protocols (Ventura et al., 1993). There were three assessment points: baseline, 1-month, and 3-months. At the baseline and 1-month assessment points, two CGI-CogS raters interviewed a patient and caregiver pair. Having two raters at two time points allowed for the calculation of inter-rater reliability for those assessments. At the 3-month assessment point, only the CGI-CogS was administered to a patient and caregiver pair and only by a single rater.

## 3. Results

### 3.1. The CGI-CogS interview

The mean time in minutes required to administer the CGI-CogS was in line with expectations for the patients (mean = 32.5,  $\pm$ 6.0), for caregivers (mean = 33.8,  $\pm$ 5.6), and the range for either type was 20 to 60 min.

#### 3.1.1. Internal consistency

Internal consistency for the CGI-CogS was evaluated using Cronbach’s alpha coefficient, which was quite high overall (see Table 2). The values are reported separately for Instrumental Functioning (IF), Social Functioning (SF), Neurocognitive State-Category Severity (NS-CS; 21 items), and for all items of the CGI-CogS (38 items). In general, these reliability statistics support the construct validity of the ADL domains and CGI-CogS. Only the Social Functioning Domain had coefficient alpha less than .88, and even this domain had alpha of .77 when considering caregiver or .76 for overall rater scores (the only alpha less than .76 was rating of the Social Functioning based exclusively on patient report).

#### 3.1.2. Inter-rater reliability

Inter-rater reliability was calculated using the intra-class correlation (ICC) between raters separately for the

Table 2  
Cronbach’s alpha coefficient for Instrumental Functioning (IF), Social Functioning (SF), and Global Severity of Cognitive Impairment Ratings (NC-CS) and total CGI-CogS scale at baseline

	Patient	Caregiver	Rater
IF (11 items)	.88	.88	.88
SF (6 items)	.74	.77	.76
NS-CS (21 items)	.90	.95	.94
CGI-CogS (38 items)	.93	.96	.95

Table 3

Inter-rater reliabilities (ICC) for Instrumental Functioning (IF), Social Functioning (SF), Global Severity of Cognitive Impairment Ratings (NS-CS and GS-CI) and Global Assessments of Cognitive Function (GAF-CogS)

CGI-CogS rating	Baseline			One month		
	Patient	Caregiver	Rater	Patient	Caregiver	Rater
IF	0.55	0.74	0.73	0.40	0.84	0.74
SF	0.43	0.65	0.75	0.47	0.65	0.61
NS-CS	0.45	0.87	0.82	0.44	0.71	0.70
GS-CI	0.60	0.80	0.71	0.50	0.82	0.80
GAF-CogS	0.58	0.63	0.60	0.66	0.67	0.68

baseline and one-month follow-up ratings. Unless noted otherwise, we used the ‘single rater’ version (ICC) to estimate the reliability of the measure if a single rating had been performed. We report the ICC’s for the Instrumental Functioning (IF), Social Functioning (SF), and the three primary global ratings of cognitive functioning, Neurocognitive State-Category Severity (NS-CS), the Global Severity of Cognitive Impairment (GS-CI), and the Global Assessment of Function for Cognition in Schizophrenia (GAF-CogS; see Table 3).

These calculations show generally acceptable levels of inter-rater reliability for the caregiver ratings, and slightly higher levels of inter-rater reliability for rater global scores (with ICC’s  $\leq .60$ ) while ratings based on patient interview alone have lower reliability (with ICC’s  $\geq .45$ ). The NS-CS ICC scores for caregiver and rater were excellent, as compared to the patient ICC which was low (ICC = .45). The ratings for the individual cognitive domains and individual items varied, with some showing acceptable to high levels of agreement, and others having lower reliability. In several instances, the reliability of the patient-based rating appeared particularly low and in those instances, the rater’s global rating also appeared to be low.

### 3.1.3. Test–retest reliability and stability

Test–retest reliability was assessed from baseline to the one-month follow-up session for three CGI-CogS variables: the Activities of Daily Living rating (ADL), the Neurocognitive State-Category Severity (NS-CS) score, and the sum of all CGI-CogS items. For these analyses, we used the mean of two raters’ global ratings at each time point. The test–retest reliability was high (see Table 4).

### 3.1.4. Description of concurrent validity analyses: correlations between CGI-CogS and Neurocognition, Activities of Daily Living, and Psychiatric Symptoms

To assess concurrent and divergent validity, the cross-sectional correlations of the CGI-CogS, Patient,

Table 4

Test–retest reliability and stability (ICC) from baseline to one-month ( $n = 33$ )

CGI-CogS score	ICC
ADL category	.92
NS-CS category	.93
GS-CI category	.95

Caregiver, and Rater globals were examined at baseline with external validity variables that included: 1) Cogtest computerized neurocognitive assessment, 2) Activities of Daily Living which included ratings of Instrumental Functioning and ratings of Social Functioning, and 3) Psychiatric symptoms (see Tables 5, 6, and 7). Because the goal of these analyses was to determine as accurately as possible the relationship between CGI-CogS ratings and these external validity variables, we used the mean of both raters’ CGI-CogS assessment.

The key objective cognitive performance variable was defined as a single Neurocognitive State Composite Score (NCS). Correlations between the Neurocognitive State-Category Severity (NS-CS) and the CGI-CogS Global Severity-Cognitive Impairment (GS-CI) scores and the NCS were in the hypothesized direction and significantly, and low to moderately correlated at baseline for the caregiver and rater global (see Table 5). Similar results were found when comparing the Global Assessment of Cognitive Functioning (GAF-CogS) and activities of daily living. Correlations between NCS and CGI-CogS ratings were generally moderate, with slightly lower values noted for ratings based solely on patient input.

We examined the correlations between the CGI-CogS Neurocognitive State-Category Severity (NS-CS) and the Global Severity-Cognitive Impairment (GS-CI) scores to Activities of Daily Living (ADL) which included the subscales of Instrumental Functioning (IF) and Social

Table 5

Correlations of CGI-CogS global ratings at baseline with neurocognitive assessment and functional outcome ( $n = 33$ )

CGI-CogS global ratings	Neurocognitive Composite Score (NCS)	Activities of Daily Living	
		Instrumental Functioning	Social Functioning
NS-CS Patient	-.14	.41*	.52**
NS-CS Caregiver	-.39*	.65**	.53**
NS-CS Rater	-.27	.57**	.51**
GS-CI Patient	-.30†	.57**	.51**
GS-CI Caregiver	-.48**	.72**	.63**
GS-CI Rater	-.41*	.69**	.57**
GAF-CogS Patient	.37*	-.58**	-.53**
GAF-CogS Caregiver	.46**	-.70**	-.62**
GAF-CogS Rater	.38*	-.68**	-.55**

† =  $p < .10$ , \* =  $p < .05$ , \*\* =  $p < .01$ .

Table 6  
Correlation of Neurocognitive Composite Score with Instrumental Functioning (IF) and Social Functioning (SF) at baseline ( $n=33$ )

	Instrumental Functioning			Social Functioning		
	PT	CR	Rater	Pt	CR	Rater
Neurocognitive Composite (NCS)	-0.37*	-0.42*	-0.41*	-0.16	-0.21	-0.18

†= $p<.10$ , \*= $p<.05$ , \*\*= $p<.01$ .

Functioning (SF). The correlations were quite high between the CGI-CogS GS-CI rater global for Instrumental Functioning and for Social Functioning. Similar results were obtained comparing the Global Assessment of Cognitive Functioning (GAF-CogS) and Activities of Daily Living (IF and SF subscales) for patient, caregiver, and rater global. We were interested in knowing how the CGI-CogS compared with objective neurocognitive functioning (Cogtest) for the prediction of Instrumental and Social Functioning (see Table 6). The finding of generally higher correlations between the CGI-CogS rater and functional outcome as compared to objective neurocognitive functioning and functional outcome warranted further evaluation (see Tables 5 and 6). We examined the difference in correlated correlation coefficients of the CGI-CogS rater globals as compared to objective neurocognitive testing composite score (NCS) in our sample of 33 patients (Meng et al., 1992). We found a significantly higher correlation for CGI-CogS GS-CI rater global ( $r=.57$ ) as compared to Cogtest ( $r=.18$ ) in predicting Social Functioning ( $t$ =test between correlated coefficients,  $t=2.39$ ,  $df=30$ ,  $p<.02$ ). There was a non-significant trend toward a higher correlation for CGI-CogS GS-CI rater global ( $r=.69$ ) as compared to Cogtest ( $r=.41$ ) in predicting Instrumental Functioning ( $t=1.98$ ,  $df=30$ ,  $p=.06$ ). There was a non-significant trend toward a higher correlation for CGI-CogS NS-CS rater global ( $r=.51$ ) as compared to Cogtest ( $r=.18$ ) in predicting Social Functioning ( $t=1.74$ ,  $df=30$ ,  $p=.09$ ). For CGI-CogS NS-CS rater global ( $r=.57$ ) compared to Cogtest ( $r=.41$ ) predicting Instrumental Functioning, the difference was not statistically significant ( $t=0.93$ ,  $df=30$ ,  $p=.37$ ). These latter comparisons showed similar trends with higher absolute values of  $r$  between CGI-CogS and measures of everyday functioning, compared to the  $r$  values relating cognitive test performance to functioning, but did not reach the threshold for statistical significance.

CGI-CogS ratings of cognition were significantly correlated with symptoms (see Table 7). The CGI-CogS Neurocognitive State-Category Severity (NS-CS) and Global Severity-Cognitive Impairment (GS-CI) scores were significantly correlated with a number of key

symptom indices such as the PANSS positive factor, PANSS negative factor, and SANS negative symptoms (see Table 6). However, cognitive functioning was not significantly correlated with BPRS depression. The results were similar when we correlated the CGI-CogS Global Assessment of Cognitive Functioning (GAF-CogS) with symptoms.

### 3.1.5. Evidence of non-linear relationships between patient report and level of cognitive deficit

While there was usually good agreement between patient and caregiver ratings, we were interested in determining whether the differences might be related to cognitive ability, given prior research suggesting that individuals with lowest levels of ability may have the least insight into their own dysfunction (Dunning, 2006; Caputo and Dunning, 2005). We examined scatterplots and used curve-fitting methods to determine the relationship between patient-caregiver discrepancies and patient performance on the Cogtest Neurocognitive Composite Score (NCS). We found a significant relationship (including both linear and quadratic components) between NCS scores and the patient-caregiver discrepancies (see Fig. 3). The results suggest that the higher the cognitive functioning of the patient, the more likely was the patient-based rating of cognitive impairment to be more severe than the caregiver-based rating. Conversely, for patients with lower cognitive functioning, the caregiver-based ratings indicated greater impairment compared to the patient-based ratings.

## 4. Discussion

We found that a new interview-based measure of cognitive function, the CGI-CogS, showed high construct validity (as indicated by Cronbach's alpha) along with adequate inter-rater reliability and good test-retest

Table 7  
Correlations of CGI-CogS ratings of Cognitive Functioning and symptoms at baseline ( $n=35$ )

CGI-CogS ratings	PANSS	PANSS	SANS	BPRS
	Pos	Neg	Total	Dep
NS-CS Patient	0.26	0.44**	0.47**	0.16
NS-CS Caregiver	0.42*	0.43**	0.45**	0.12
NS-CS Rater	0.36*	0.42*	0.41*	0.13
GS-CI Patient	0.41*	0.54**	0.61**	0.09
GS-CI Caregiver	0.48*	0.45**	0.52**	0.18
GS-CI Rater	0.37*	0.42*	0.49**	0.12
GAF-CogS Patient	-0.42*	-0.50**	-0.62**	0.01
GAF-CogS Caregiver	-0.50**	-0.51**	-0.63**	-0.12
GAF-CogS Rater	-0.49**	-0.49**	-0.61**	-0.11

†= $p<.10$ , \*= $p<.05$ , \*\*= $p<.01$ .

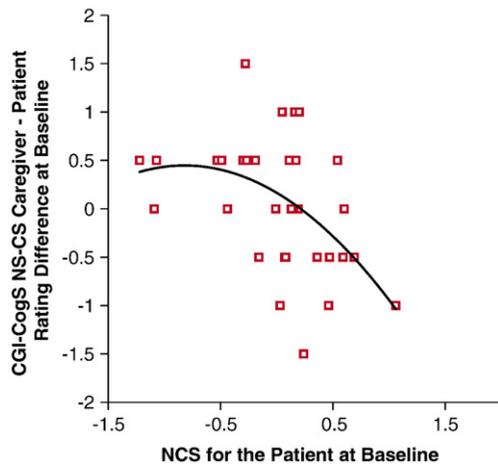


Fig. 3. Scatter plot showing Baseline Cogtest NCS Patient Scores (x axis) and Baseline Caregiver–Patient NS-CS Rating Difference Scores (y axis). The curve fits a quadratic function to these data, accounting for 5% more variance than the linear trend.

reliability over a one month interval, supporting its utility in the assessment of people with schizophrenia. We generally found moderate correlations between CGI-CogS ratings and both objective neurocognitive test results and day-to-day instrumental and Social Functioning. In fact, the CGI-CogS ratings were significantly more correlated with Social Functioning than were objective neurocognitive test scores ( $p=.02$ ) and showed trends in the same direction for predicting Instrumental Functioning ( $p=.06$ ), despite modest sample sizes. Therefore, the CGI-CogS appears to satisfy several of the important reliability and validity criteria suggested by the MATRICS project for interview based, co-primary measures of cognition, and there is evidence suggesting that CGI-CogS ratings are more correlated with functional outcomes than are objective performance-based tests of neurocognition.

Our results were in several ways similar to those of Keefe et al. (2006) who also studied an interview-based measure of cognition, the Schizophrenia Cognition Rating Scale (SCoRS). Among the similarities, both studies suggest that ratings based on patient report alone may have lower validity than ratings based on caregiver input, or lower than the combined information obtained from both patient and caregiver that is then integrated by a trained rater. One difference between the CGI-CogS and SCoRS may be reflected by the internal consistency of the instruments (coefficient alpha for CGI-CogS Neurocognitive State category = .94; coefficient alpha for SCoRS total = .79). This is likely due to higher average item-total correlations, perhaps because each CGI-CogS item includes multiple probes (i.e., CGI-

CogS is more like an ‘interview’ while SCoRS is more like a ‘questionnaire’). This difference in the construction of the instruments may explain why we tended to see greater evidence for validity even for the patient-only report, while the SCoRS ratings based only on patient report were clearly lower than those based on caregiver and global ratings based on all sources of information. This difference in internal consistency may also have implications for identifying appropriate short-forms of these instruments. Despite these differences, the results from both our study and Keefe et al. (2006) suggest the value of obtaining caregiver input in addition to examining the patient alone. Not only did ratings based on caregiver input tend to have higher validity with respect to other measures, but the trained raters’ final ratings from all sources tended to be influenced more often by caregiver input.

An interesting additional finding from our study may help shed light on the nature of discrepancies between patient and caregiver reports, and also why prior research has found relatively poor agreement between schizophrenia patients’ self-report ratings of neurocognition and their objectively assessed neurocognitive performance (Prouteau et al., 2004; Moritz et al., 2004; Jungwirth et al., 2004). We have shown evidence of a non-linear relationship which indicates that patients with lower levels of objectively-assessed cognitive competency appeared to have a correspondingly lower ability to provide valid information about their cognitive function. More specifically, patients with lower levels of cognitive competency tend to overestimate their abilities, paralleling extensive research on self-evaluation in non-psychiatric groups (Dunning, 2006; Caputo and Dunning, 2005). Less expected was the observation that patients with higher levels of objectively assessed cognitive ability appeared to underestimate their own abilities relative to caregiver reports. This negative bias is similar to findings in clinical samples of individuals with major depressive disorder (Beck, 1976; Miller and Norman, 1986; Cane and Gotlib, 1985; Mogg et al., 2006). Additional research might fruitfully examine further the sources of discrepancy between patient report and other sources of information. Until such research is conducted, our findings support the conclusion that interview-based measures of cognitive function in schizophrenia benefit from including multiple sources of information. At the same time, it is important for research and clinical practice not to minimize the value of patient’s impressions about their own cognitive functioning, as these may be the most critical factors in engaging patients productively in therapeutic relationships and rehabilitation efforts (Prouteau et al., 2004).

Because of the potential for overlap of symptoms and cognitive complaints, the study raters were instructed to distinguish symptom influences on functional outcomes and make their ratings based on the cognitive deficits alone. The CGI-CogS ratings show little overlap with depressive symptoms, but there was considerable overlap with ratings of negative symptoms, and with positive symptoms. The negative symptom overlap was to be expected given the widely observed cross-sectional correlations of negative symptoms with cognitive deficits (Harvey et al., 2006). The positive symptom overlap was more surprising, given that there are usually no clear relationships between positive symptoms and cognitive performance deficits. This suggests that positive symptoms might influence interview-based ratings of cognition. However, the correlations between objective cognitive measures and positive symptoms may reflect our design, which focused on inclusion of clinically stable patients with principally persistent positive symptoms. The FDA has expressed concerns about “pseudospecificity” in the context of identifying new indications such as cognitive impairment associated with schizophrenia (i.e., applicants for this new indication may be required to demonstrate that cognitive deficits are not somehow “secondary” to symptoms traditionally used to define the syndrome). The observed correlations may arise because the CGI-CogS and symptom rating scales both rely on patient report, and thus share method variance. In addition, both positive symptom scales and the CGI-CogS are assessing the past month, while objective cognitive assessments are usually conducted at one point in time. This overlap in source of information (patient) and time period covered could have led the positive symptoms and CGI-CogS ratings to be more highly correlated with each other, and reduced the correlation of positive symptoms with the Cogtest scores. In addition, there could be a relationship between the patient’s perception of his or her own cognitive deficits, which influence the CGI-CogS ratings, and positive symptom reports that also require the patient to acknowledge a problem. All these factors increase the likelihood of finding a relationship even if positive symptoms and neurocognitive deficits are not truly related. An empirical question remains whether the correlation of positive symptoms with CGI-CogS ratings is better explained as an “artifact” of positive symptoms or a real-life consequence of certain cognitive impairments that are not yet well measured by objective cognitive tests.

Interview-based measures of cognition potentially have several theoretical and practical advantages over objective neurocognitive tests that could ultimately facilitate their broader applicability in clinical research and clinical practice. Compared to objective tests, interview-based

measures are shorter and therefore require less time to administer, score, and interpret. Interview-based measures may be more amenable to use by clinicians who do not possess specific training in psychometrics, neurocognitive testing, or measurement of intelligence. Raters that administer interview-based measures of cognition, assuming they have good clinical skills, could more easily be trained to assess cognition through didactic instruction and ratings of “gold standard” training videotapes, and standards exist enabling objective rater certification. With interview-based measures, rater fidelity checks also can be conducted, increasing rater consistency even in multi-site studies and within raters over time.

There might be several important uses for interview-based measures of cognition beyond their value as endpoints in clinical trials. Interview-based measures of cognition could serve to augment existing clinical evaluation procedures when attempting to identify and engage in treatment individuals who are experiencing cognitive deficits that impair functioning. Since the CGI-CogS clearly targets issues of concern to the patients and their caregivers, it has both more obvious relevance to them, and highlights for treating clinicians those areas that are important from their perspective. These features suggest that interview-based measures might foster enhanced communication and stronger treatment alliances that are important in augmenting the efficacy of both pharmacological and non-pharmacological interventions. Additional benefits could be realized because the format of CGI-CogS includes 7 cognitive domain headings and comprehensible definitions of basic neurocognitive functions. The CGI-CogS might therefore help disseminate knowledge about cognitive functions, and their relationship to functional outcomes, which may be valuable for psychiatrists, caregivers, case managers and other people interested in promoting the health and well being of people with schizophrenia.

The current study has several limitations. The possibility that the caregiver ratings were biased by patient reports, because the patients were usually interviewed before the caregiver, cannot be ruled out. For example, the higher validity of caregiver information might only be observed when the patient is interviewed first. In addition, the results may be influenced by the use of case managers as caregivers as compared to the use of other caregivers, such as family members or friends. The case managers who participated in the study might have provided especially valid insight into the cognitive functioning of the patients by virtue of their greater training and experience. Indeed these “caregivers” have worked in a research environment where neurocognitive studies are conducted routinely. The correlations between

the CGI-CogS ratings and daily functioning also might have been more robust because of shared method variance and rater overlap. Finally, our patients were selected specifically for clinical stability and thus we cannot comment on the sensitivity of CGI-CogS to real cognitive change that may be associated with changes that result from effective treatment. This issue, of course, awaits development of a treatment that can accomplish this goal.

Future directions for research on the CGI-CogS as a co-primary measure of cognition include: (1) refinement of specific item content; (2) application in intervention studies to better document reliability and sensitivity to change; (3) explicit separation of patient–caregiver input to better assess the possible utility of a caregiver-only instrument, or separate instruments for patients and caregivers; and (4) use of classical and modern psychometric methods, including Item Response Theory, to further refine and shorten the instrument; these plans are now underway.

#### Role of the Funding Source

Funding for this study was provided by a grant to Dr. Robert Bilder from Pfizer Inc. Pfizer Inc. had no further role in the study design, collection, analysis, interpretation of data, writing of the report, or the decision by the authors to submit the paper for publication.

#### Contributors

Dr. Bilder wrote the protocol, supervised the conduct of the study, trained the raters in conducting the interview-based measures of cognition, conducted the majority of the statistical analyses, and commented on early drafts of the paper. Dr. Ventura assisted Dr. Bilder with all aspects of the study and trained the raters on diagnostic and symptom assessment, recruited study participants, conducted cognitive and symptom assessments, conducted data analyses, performed literature searches, and drafted the manuscript. Dr. Cienfuegos and Mr. Boxer recruited study participants, conducted cognitive and symptom assessments, and commented on drafts of the paper. Mr. Boxer entered data and managed study data files. All authors have contributed to and have approved the final manuscript.

#### Conflict of Interest

This study titled, Clinical Global Impression of Cognition in Schizophrenia (CGI-CogS): Reliability and Validity of a Co-Primary Measure of Cognition, was supported by a grant to Dr. Robert Bilder from Pfizer Inc. In addition, Dr. Bilder has acted as a consultant for Pfizer Inc. Dr. Joseph Ventura has received a grant from Pfizer Inc. and from the NIMH to conduct further scale development on the CGI-CogS that was initiated in the current study. Drs. Bilder, Ventura, and Cienfuegos received royalties for use of the CGI-CogS.

#### Acknowledgment

This paper was presented in part at the 61th Annual Meeting of the Society of Biological Psychiatry, May 18–20, 2006, Toronto, Canada. The authors wish to thank the study participants and the following colleagues who supported the project at the UCLA Department of Psychiatry: Keith Nuechterlein, Ph.D., and Barry Guze, M.D., and the

Department of Veterans Affairs VISN 22 Mental Illness Research, Education, Clinical Center: Steven Marder, M.D.

#### References

- Andreasen, N.C., 1984. The Scale for the Assessment of Negative Symptoms (SANS). The University of Iowa, Iowa City, IA.
- Beck, A.T., 1976. Cognitive therapy and the emotional disorders. International University Press, New York.
- Bilder, R., Ventura, J., Cienfuegos, A., 2003. Clinical Global Impression of Cognition in Schizophrenia (CGI-CogS) Manual and Rating Sheet. UCLA Department of Psychiatry, Los Angeles, California.
- Cane, D.B., Gotlib, I.H., 1985. Depression and the effects of positive and negative feedback on expectations, evaluations, and performance. *Cognitive Therapy and Research* 9 (2), 145–160.
- Caputo, D., Dunning, D., 2005. What you don't know: the role played by errors of omission in imperfect self-assessments. *Journal of Experimental Social Psychology* 41, 488–505.
- Carpenter, W., 2004. Clinical constructs and therapeutic discovery. *Schizophrenia Research* 72, 69–73.
- Dunning, D., 2006. Strangers to ourselves? *The Psychologist* 19, 600–603.
- Gold, J., 2004. Cognitive deficits as treatment targets in schizophrenia. *Schizophrenia Research* 72, 21–28.
- Green, M.F., 1996. What are the functional consequences of neurocognitive deficits in schizophrenia? *American Journal of Psychiatry* 153, 321–330.
- Green, M., Nuechterlein, K.H., 1999. Should schizophrenia be treated as a neurocognitive disorder? *Schizophrenia Bulletin* 25, 309–318.
- Green, M., Nuechterlein, K.H., 2004. The MATRICS initiative: developing a consensus cognitive battery for clinical trials. *Schizophrenia Research* 72, 1–3.
- Green, M.F., Kern, R.S., Braff, D.L., Mintz, J., 2000. Neurocognitive deficits and functional outcome in schizophrenia: are we measuring the “right stuff”? *Schizophrenia Bulletin* 26, 119–136.
- Green, M.F., Kern, R.S., Heaton, R.K., 2004. Longitudinal Studies of Cognition and Functional Outcome in Schizophrenia: Implications for MATRICS. *Schizophrenia Research* 72, 41–51.
- Guidance for Industry Patient-reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims Draft Guidance. U.S. Department of Health and Human Services Food and Drug Administration, February, 2006, Washington, D.C.
- Harvey, P.D., Serper, M.R., White, L., Parrella, M.J., McGurk, S.R., Moriarty, P.J., Bowie, C., Vadhan, N., Friedman, J., Davis, K.L., 2001. The convergence of neuropsychological testing and clinical ratings of cognitive impairment in patients with schizophrenia. *Comprehensive Psychiatry* 42, 306–313.
- Harvey, P.D., Koren, D., Reichenberg, A., Bowie, C.R., 2006. Negative symptoms and cognitive deficits: what is the nature of their relationship? *Schizophrenia Bulletin* 32, 250–258.
- Jungwirth, S., Fischer, P., Weissgram, S., Kirchmeyr, W., Bauer, P., Tragl, K., 2004. Subjective memory complaints and objective memory impairment in the vienna-transdanube aging community. *Journal of the American Geriatrics Society* 52, 263–268.
- Kay, S., Fiszbein, A., Opler, L., 1987. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin* 13, 261–276.
- Keefe, R., Poe, M., Walker, T., Kang, J., Harvey, P., 2006. The schizophrenia cognition rating scale: an interview-based assessment and its relationship to cognition, real-world functioning, and functional capacity. *American Journal of Psychiatry* 163, 426–432.

- Meng, X.-L., Rosenthal, R., Rubin, D.B., 1992. Comparing correlated correlation coefficients. *Psychological Bulletin* 111, 172–175.
- Miller, I.W., Norman, W.H., 1986. Persistence of depressive cognitions within a subgroup of depressed inpatients. *Cognitive Therapy and Research* 10 (2), 211–224.
- Mogg, K., Bradbury, K.E., Bradley, B.P., 2006. Interpretation of ambiguous information in clinical depression. *Behaviour Research and Therapy* 44 (10), 1411–1419.
- Moritz, S., Ferahli, S., Naber, D., 2004. 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* 10, 623–633.
- Nuechterlein, K.H., Miklowitz, D.J., Ventura, J., Gitlin, M.J., Stoddard, M., Lukoff, D., 2006a. Classifying episodes in schizophrenia and bipolar disorder: criteria for relapse and remission applied to recent-onset samples. *Psychiatry Research* 144, 153–166.
- Nuechterlein, K.H., Barch, D.M., Gold, J.M., Goldberg, T.E., Green, M.F., Heaton, R.K., 2006b. Identification of separable cognitive factors in schizophrenia. *Schizophrenia Research* 72, 29–39.
- Prouteau, A., Verdoux, H., Briand, C., Lesage, A., Lalonde, P., Nicole, L., Reinhartz, D.A., Stip, E., 2004. Self-assessed cognitive dysfunction and objective performance in outpatients with schizophrenia participating in a rehabilitation program. *Schizophrenia Research* 69, 85–91.
- Silverstein, S., Wilkniss, S., 2004. At issue: the future of cognitive rehabilitation of schizophrenia. *Schizophrenia Bulletin* 30, 679–692.
- van den Bosch, R.J., Rombouts, R.P., 1998. Causal mechanisms of subjective cognitive dysfunction in schizophrenic and depressed patients. *Journal of Nervous and Mental Disease* 186, 364–368.
- Ventura, J., Green, M., Shaner, A., Liberman, R.P., 1993. Training and quality assurance on the BPRS: “The Drift Busters”. *International Journal of Methods in Psychiatric Research* 3, 221–244.
- Ventura, J., Nuechterlein, K.H., Subotnik, K.L., Gutkind, D., Gilbert, E.A., 2000. Symptom dimensions in Recent-Onset Schizophrenia: and Mania: a principal components analysis of the 24-item Brief Psychiatric Rating Scale. *Psychiatry Research* 97, 129–135.