Assessing cognition and real-world functioning in schizophrenia

Summary

With the new focus on functional recovery in schizophrenia, factors limiting function in schizophrenia are receiving increasing attention. Neurocognitive (NC) impairment accounts for 20-60% of the variance in real-world outcome. The effect sizes of the associations between NC and functional outcome tend to be medium for specific domains and larger for summary scores. Mapping NC deficits often requires the use of extensive test batteries that are lengthy and costly and require advanced training in assessment to score and utilize the results. The currently available NC assessment instruments differ widely in the population intended for use, administration time, interpretation of results, and the assessment of certain NC domains.

Social cognition (SC) contributes to functional outcomes beyond the influence of NC and may have a greater impact than NC on social outcomes. In addition, SC may mediate the relationship between NC and social functions in both chronic and first-episode patients. The degree of the relationship between SC and functioning varies, depending on the SC domain and the type of functional outcome assessed. In the past, there has been controversy over what SC processes should cover. Moreover, the most critical issue is that there is no consensus in the field as to which measures best assess each SC domain. As a result, a heterogeneous group of tasks have been administered with significant conceptual overlap and questionable psychometric properties across studies. This problem is present across all SC domains and contributes to the inconsistency of the reported findings.

The assessment of real-life functioning in schizophrenia presents complex challenges from variability in the operational definition of functional outcome to problems in identifying optimal information sources. In this context, there are still few satisfactorily reliable instruments for the assessment of functional outcomes that are practical in terms of time involved, and most real-life functional outcome scales seem to be largely redundant with each other when utilized simultaneously.

This update describes the main NC and social cognition (SC) batteries and real-world assessments used in schizophrenia and discuss their advantages and disadvantages.

Key words

Schizophrenia • Neurocognition • Social cognition • Real-world functioning • Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) • Specific Levels of Functioning (SLOF)

Introduction

During the last two decades, the field of psychiatry has moved toward the goals of remission and recovery rather than mere symptom improvement 1. Recovery refers to patients being able to function normally at work or school, in the community, and at home. It may occur even if patients are still experiencing some ongoing symptoms. Recovery can also involve a host of subjective experiences including attaining a self-appraised acceptable quality of life or reasonable sense of social rank and recapturing a cohesive sense of oneself as a valuable person in the world. Recovery requires that the person diagnosed with schizophrenia be an active agent in that process 2. It has been suggested that only 13.7% of subjects in their first episode of schizophrenia or schizoaffective disorder met full recovery criteria for 2 years or longer 3.
A key focus of current research is to identify factors limiting function in schizophrenia. Neurocognitive (NC) impairment has long been recognized as a core component of schizophrenia and is closely linked to social and occupational outcome. This association between NC and outcome is robust—it was replicated and extended in many countries, using many different types of assessments, in different patient groups across phase of illness, including prodromal. NC impairment accounts for 20-60% of the variance in real-world outcome and has been shown to predict social outcomes more closely than do psychotic symptoms. However, effect sizes of the associations between NC and functional outcome tend to be medium for specific domains and larger for summary scores. Thus, the questions have shifted from whether NC is related to outcome to how NC is related to outcome. Further, not all types of NC are equally important when it comes to navigating the real world. In a 1996 review, Green has shown that several NC domains were associated with specific functional outcomes in schizophrenia. The strongest evidence showed that verbal memory was associated with all measures of functional outcomes; moreover other correlations were found between vigilance and both social problem solving and skill acquisition, and between card sorting and community functioning. This growing realization that NC deficits are central to outcomes have directed attention at the assessment of these important aspects of the disorders. This update describes the main NC and social cognition (SC) batteries and real-world assessments used in schizophrenia and discuss their advantages and disadvantages.

Cognitive assessment
NC deficits often precede the manifestation of psychosis and might be the first signs of schizophrenia in at-risk patients, they are orthogonal to positive and negative symptoms, are relatively stable over time, continue to be present after remission of psychosis, and are relatively unaffected by antipsychotic treatment. Moreover, although NC assessment does not represent a clear diagnostic marker, as a clear profile on neuropsychological tests has not been defined, it can offer an understanding of NC deficits and guide treatment targets and recommendations. Several investigators have emphasized different area of cognition. The Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) initiative, sponsored by the United States National Institute of Mental Health (NIMH), identified seven distinct, separable cognitive domains as commonly deficient in schizophrenia: attention/vigilance, working memory, reasoning and problem solving, processing speed, visual learning and memory, verbal learning and memory, and social cognition. The seventh domain, SC, was included because it was viewed as an ecologically important domain of cognitive deficit in schizophrenia that shows promise as a mediator of NC effects on functional outcome. Some NC domains were not included in this list, i.e., verbal comprehension was not included in the cognitive battery as it was considered resistant to change.

A growing body of the literature suggests that many of these deficits can be traced to a generalized cognitive impairment. Patients with schizophrenia have been shown to perform below the level of their peers by over one standard deviation (SD) and have a unique profile of impairment of cognitive domains.

Thus, the evaluation of a patient with schizophrenia should begin with a general assessment to evaluate the patient’s average cognitive functioning, investigating the school or works performances, any developmental delays or change in overall functioning, and the ability to solve problems in daily lives. Proverb interpretation brings out unusual thought content and deficits in executive function.

Tests for measuring cognition
Many instruments are available for the assessment of cognitive functioning in patients with schizophrenia. These instruments differ widely in the population intended for use, administration time, interpretation of results, and the assessment of certain cognitive domains, and little guidance is available for selection among these instruments for clinical trials. Mapping cognitive deficits often requires the use of extensive test batteries that are lengthy and costly and require advanced training in assessment to score and utilize the results. In particular, due to the nature of deficits in schizophrenia, it is reasonable to suggest the need for an affordable, easy to administer test that identifies deficits in cognitive skills in order to recommend an intervention for addressing these deficits. Useful tests for clinical practice are described below.

Wechsler Adult Intelligence Scale
Abbreviated versions of the Wechsler Adult Intelligence Scale (WAIS) have been developed as pragmatic timesaving devices that balance the length of assessment with accurate estimates of the overall level of general intellectual functioning. Two main types of abbreviations of the WAIS have been proposed: “select-subtest” and “select-item”. Selected subtests abbreviations reduce the amount of time spent on test assessment by only administering selected subtests to obtain estimated cognitive functioning scores. This selection can vary from seven to as few as two subtests, in which,
for example, only the subtests Vocabulary and Block Design are being administered. Select-item abbreviations involve the administration of previously selected items from all subtests. The Satz-Mogel short form \(^{19}\) is frequently used as a “select-item” abbreviation and has shown to be an accurate measure of general intellectual ability when compared with the full WAIS \(^{20}\). In the Satz-Mogel method, the item selection concerns the administration of every second (or third) item in WAIS subtests (e.g., Information, Block Design) that take a long time to administer.

The proposed 15-minute version of the WAIS by Verlhorst and colleagues \(^{21}\), that includes only select items from three subtests, may serve as a useful screening device for general intellectual ability in research or clinical settings, and is recommended when a quick and accurate IQ estimate is desired.

**Repeatable Battery for the Assessment of Neuropsychological status**

The Repeatable Battery for the Assessment of Neuropsychological status (RBANS) is a standardized screening instrument designed to assess global neuropsychological functioning in a brief administration. Several studies have supported the psychometric properties of the RBANS \(^{22}\), with past research reporting acceptable test-retest reliability, internal consistency, and concurrent validity \(^{23}\). The RBANS has been found to be a valid measure of the cognitive decline associated with various neurological conditions including stroke, Alzheimer's disease, multiple sclerosis, Parkinson's disease, and Huntington's disease. This instrument measures several cognitive domains of interest in schizophrenia – immediate memory, visuospatial/constructional ability, language, attention, and delayed memory – and provides a global measure, the total scale score. In addition, the RBANS offers two alternate forms to reduce the potential influence of practice effects in serial test administration. Wilk and colleagues \(^{24}\) found that RBANS was reliable and sensitive enough to differentiate between patients with schizophrenia and healthy subjects and was suited for repeated measures.

**Measurement and Treatment Research to Improve Cognition in Schizophrenia Consensus Cognitive Battery**

The Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) Consensus Cognitive Battery (MCCB) \(^{25}\) was designed by the NIMH to support the development of pharmacological agents for improving the neurocognitive impairments in schizophrenia. It has been recommended by the United States Food and Drug Administration (FDA) to assess cognitive impairment as the primary outcome measure in registry trials of schizophrenia \(^{26}\). An initial MATRICS consensus conference involving more than 130 scientists from academia, government, and the pharmaceutical industry led to agreement on seven cognitive domains for the battery and on five criteria for test selection. The criteria emphasized characteristics required for cognitive measures in the context of clinical trials: test-retest reliability; utility as a repeated measure; relationship to functional status; potential changeability in response to pharmacological agents; and practicality for clinical trials and tolerability for patients. Cognitive function was measured according to the 7 cognitive domains of the MCCB derived from scores on 10 cognitive measures: speed of processing (Trail Making Test Part A; Brief Assessment of Cognition in Schizophrenia: Symbol coding; Category fluency test, animal naming), attention/vigilance (Continuous Performance Test: Identical Pairs), working memory (Wechsler Memory Scale, spatial span subset; Letter Number Span test), verbal learning (refers to immediate verbal memory, Hopkins Verbal Learning Test (HVLT)-Revised, immediate recall), visual learning (refers to immediate visual memory, Brief Visuospatial Memory Test-Revised), reasoning and problem solving (Neuropsychological Assessment Battery (NAB), mazes subtest), and social cognition (Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT); managing emotions branch). The MCCB can be administered in 1 to 1.5 hours.

After substantial use in the testing ground of multisite clinical trials, the MCCB has demonstrated impressive psychometrics. It has shown sensitivity to improvement from interventions, most notably for cognitive training interventions so far. The MCCB also tracks with key biomarkers, which is an important feature as efforts are made to apply experimental medicine principles, such as target engagement, to psychiatric treatment trials. Two limitations of the MCCB should be pointed out. First, the MCCB was developed to facilitate drug approval from the U.S. FDA. Because clinical trials are often international, it soon became obvious that a key limitation of the MCCB was that it was available only in English. Hence, the MCCB has now been professionally translated and is commercially available in over 20 languages (see www.matricsinc.org for a listing). With the help of an industry-academic-government consortium (MATRICS-CT), representative normative data were collected in key countries and these norms were used to create international scoring programs. Recently, Mucci et al. \(^{26}\) have reported the normative Italian data. A second change reflects the growing awareness that NC and SC are separable dimensions. It is possible that a treatment will affect SC and NC differently, and that is assumed for specific interventions, that are focused on one area or the other. To allow trial investigators to examine NC only, the MCCB scoring program now provides an option for a “neurocognitive composite” that
Brief Assessment of Cognition in Schizophrenia

The Brief Assessment of Cognition in Schizophrenia (BACS) \(^{27}\) is a portable pen-and-paper, concise tool designed to evaluate five different domains of cognitive function with six tests (impairment of verbal memory, working memory, motor speed, verbal fluency, attention and processing speed, and executive function), assessing the aspects of cognition found to be most impaired and most strongly correlated with outcome in patients with schizophrenia. The BACS is easily performed in clinical settings and can be administered by medical professionals in 30-35 minutes. It yields a high completion rate in these patients, and has high reliability. The BACS was found to be as sensitive to cognitive impairment in patients with schizophrenia as a standard battery of tests that required over 2 h to administer. The way used to create the BACS and the RBANS was to create a new group of tests that assess all or many of the key domains of neuropsychological functioning in less time than the comprehensive batteries traditionally used. With this approach, one can create tests that are specifically sensitive to the deficits in schizophrenia patients and likely to be amenable to change with atypical antipsychotics. Assessment batteries such as these require the collection of large amounts of data to establish population norms, reliability, and validity. These psychometric properties make the BACS a promising tool for assessing cognition repeatedly in patients with schizophrenia, especially in clinical trials of cognitive enhancement.

Brief cognitive assessment

The Brief cognitive assessment (BCA) consists of three standard tests selected from among those commonly included in comprehensive cognitive batteries administered to patients with schizophrenia: Verbal Fluency (letters and categories), Trails A and B, and the Hopkins Verbal Learning Test. The way used to create the BCA was to select a small number of standardized tests widely used in clinical neuropsychology, that examine a variety of cognitive domains in a very limited fashion. With this approach, normative data and information regarding the sensitivity, reliability and validity of the individual tests included in the battery are already available. The creation of the BCA was guided by several principles. First, the battery had to be very brief (maximum 15 min), including time for set-up, administration, and scoring. Second, tests in the BCA had to be easy to administer and score and the results had to be understandable to clinicians who were not neuropsychologists. Third, the battery had assess aspects of each of the cognitive domains known to be impaired in patients with schizophrenia, including executive functions, memory, attention, and processing speed. Fourth, the tests selected had to have been found to be sensitive to improvements with atypical antipsychotics. Finally, the tests had to have been found to be related to measures of functional outcome in schizophrenia patients.

Schizophrenia Cognition Rating Scale

The Schizophrenia Cognition Rating Scale (SCoRS) \(^{28}\) assesses the following cognitive domains: memory, working memory, attention, reasoning and problem solving, language and motor skills. The SCoRS has several advantages including brief administration time, approximately 15 min per interview, association to real-world functioning, good test-retest reliability, and correlations with performance-based measures of cognition \(^{28}\). Moreover, its rating is based on information from three separate sources including patient, informant of the patient who has regular contact with patient and interviewer. The SCoRS has shown good reliability, validity, and sensitivity to cognitive impairment in schizophrenia, with the advantage of brief administration and scoring time. The SCoRS was developed to measure cognitive functions through questions about cognitions related to daily life events \(^{28}\). It consists of 20 items. Each item is rated on a scale ranging from 1 to 4 with higher scores reflecting a greater degree of impairment. Every item is given anchor points based on the degree of their daily problems. Two studies have demonstrated significant correlations between SCoRS ratings with NC functions as well as psychosocial functioning, in Singaporean, and Italian schizophrenia patients \(^{29,30}\).

Social cognition assessment

SC impairments may precede onset of the schizophrenic illness \(^{31-33}\) and are present early in the illness \(^{34}\). Such SC deficits have been consistently linked to a variety of real-world outcomes, such as social competence, community functioning, and quality of life \(^{35}\). The extent of overlap between SC and NC has been an area of debate within the literature. SC contributes to functional outcomes beyond the influence of NC and may have a greater impact than NC on social outcomes \(^{36-38}\). In addition, SC may mediate the relationship between NC and social functions in both chronic \(^{28,41}\) and first-episode patients \(^{42}\). Moreover, treating SC deficits leads to improvements in real-world social outcomes, including social adjustment, social functioning, social relationships, aggressive incidents, and social skills \(^{43}\). The degree of the relationship between SC and functioning varies, depending on the SC domain and the type of functional outcome assessed.
Even if the study of SC is quite robust, its study is in some ways less developed than that of NC. In the past, there has been controversy over what SC processes should cover. Moreover, the most critical issue is that there is no consensus in the field as to which measures best assess each SC domain. As a result, a heterogeneous group of tasks have been administered with significant conceptual overlap and questionable psychometric properties across studies. This problem is present across all SC domains and contributes to the inconsistency of the reported findings.

The Social Cognition Psychometric Evaluation (SCOPE) study was designed to address these challenges: first, to achieve a consensus on the crucial SC domains in schizophrenia; second, to evaluate the psychometric properties of existing measures and their suitability for clinical trials. Using methods similar to other NIMH measurement initiatives (e.g., MATRICS, MATRICS-CT, and VALERO), the panel of experts in the schizophrenia-spectrum research field identified and supported the value of four key domains, including emotion processing (EP; the ability to perceive and appropriately use emotions), theory of mind (ToM; the ability to infer one’s own and others’ mental states), social perception (SP; ability to decode and interpret social cues in others), and attributional style (AS; ability to explain the causes or make sense of social interactions and events). Two additional domains, social metacognition and social reciprocity, however suggest avenues for expansion of SC research.

Using RAND consensus ratings, the expert survey of SC produced 108 different outcomes measures, with many of these domains and measures being very closely related to each other. The panelists lastly selected the following measures for further evaluation: Ambiguous Intentions Hostility Questionnaire, Bell Lysaker Emotion Recognition Task, Penn Emotion Recognition Test, Relationships Across Domains, Reading the Mind in the Eyes Test, The Awareness of Social Inferences Test, Hinting Task, and Trustworthiness Task. However, the similarity of many of these measures to each other has led to challenges in direct comparisons of their usefulness, as many of these assessments have overlapping content. Moreover, only a limited amount of psychometric information is currently available for the candidate measures, which underscores the need for well-validated and standardized measures in this area.

In the context of a multicenter study of the Italian Network for Research on Psychoses (NIRP), the assessment of SC included a test contained in the MCCB: the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) managing emotion section, which examines the regulation of emotions in oneself and in one’s relationships with others by presenting vignettes of various situations, along with ways to cope with the emotions depicted in these vignettes. It was integrated by the Facial Emotion Identification Test (FEIT), which examines emotion perception, and The Awareness of Social Inference Test (TASIT), which is a TOM test consisting of 7 scales (positive emotions, negative emotions, sincere, simple sarcasm, paradoxical sarcasm, sarcasm enriched, lie), organized into three sections: Emotion recognition; Social Inference (minimal); Social Inference (enriched). The manual of the TASIT was translated into Italian by a psychiatrist of the Department of Psychiatry of the University of Naples SUN. The videotaped vignettes of the TASIT were dubbed in Italian at the Fono Roma Studios (www.fonoroma.com). As to the FEIT, the adaptation of the Italian version required the translations of the six emotions reported on the screen above the stimuli.

Real-world functioning assessment

The assessment of real-life functioning presents complex challenges from variability in the operational definition of functional outcome to problems in identifying optimum information sources. Indeed, many different strategies have been proposed to assess real-life functioning, including self-report interviews, proxy reports, informant interviews, direct observations by trained clinicians, and performance-based measures, which assess functional capacity (“what a person is able to do under optimal conditions”). However, reports of real-life outcomes vary across informants and contain elements of error or shortcomings. It has been suggested that self-reports should be accepted at face value even if they reflect patients’ delusional beliefs and have limitations such as inaccurate estimations. Other investigators have highlighted the potential for psychotic symptoms, mood states, disorganized thinking, lack of insight, and NC deficits to limit the usefulness of the self-report methodology in severely ill schizophrenia patients. Furthermore, it has been suggested that these measures may not adequately reflect the effects of various interventions. However, studies have shown that patient self-reports of everyday functioning in schizophrenia often do not converge with objective evidence or the reports of others. Self-reports of functioning therefore appear problematic, and alternative assessment methods may be required. However, many patients have no caregivers to provide information, and variance in their reports can be influenced by the amount of contact with the subject and situation specificity of the observation. High contact clinicians appear to generate ratings of everyday functioning that are more closely linked to patients’ ability scores.
Overview of everyday real-life outcomes

In this context, research efforts are increasingly turning to the design, evaluation and improvement of relatively economical real-life measurement 59-61. Moreover, given concerns about length and ease of administration, as well as burden to the subject for assessment batteries, a practical measure must be both cost efficient and require a modest amount of time to administer 59. However, there are still few satisfactorily reliable instruments for the assessment of functional outcomes that are practical in terms of time involved, and most real-life functional outcome scales seem to be largely redundant with each other when utilized simultaneously. One upshot of this situation is the Validation of Everyday Real-World Outcomes in schizophrenia (VALERO Expert panel) initiative. The goal of this initiative was to identify the functional rating scale or scales (or subscales from existing scales) (self-report and informant-based reports) most strongly related to performance-based measures of cognition and everyday living skills through a comprehensive evaluation of existing instruments 49. The outcomes may include social, vocational, independent living, self-care or any combination of these. The scale characteristics, which were rated by the panelists and were similar to those deemed important in the MATRICS process, were: reliability (test-retest and interrater), convergence with performance-based measures of functional capacity and neurocognitive performance, sensitivity to treatment effects, usefulness for multiple informants (e.g., self, friend or relative, case manager, or prescriber), relationships with symptom measures, practicality and tolerability for people with low education levels, and convergence with other measures of real-life functional outcomes (including either other rating scales or achievement milestones). Among the 59 measures nominated, the investigators selected the 11 scales that were the most highly nominated, had the most published validity data regarding their psychometric qualities and best represented the domains of interest (social functioning, everyday living skills, or both these areas - "hybrid" scales). Scales were rated on a 9-point (1-9) scale, where scores of 1-3 were poor, 4-6 were fair to good and 7-9 were very good to superb. The two scales that scored highest across the various criteria for each of the classes of scales (hybrid, social functioning, and everyday living skills) were selected for use in the first substudy of VALERO 49. The scales selected were the Quality-of-Life Scale, Specific Levels of Functioning Scale, Social Behavior Schedule, Social Functioning Scale, Independent Living Skills Schedule, and Life Skills Profile. The overall results of this first substudy of VALERO show that all examined scales can be considered as somewhat useful in their current versions. Moreover, many of these scales lack critical data regarding reliability across investigators and relationship with neuropsychiatric and functional capacity performance. Ratings for usefulness across multiple raters were also quite low, partly because many of these scales do not have alternate forms that attempt to capture the differing perspectives of different raters. As an entirely effective measure of the real-life outcomes component of the functional outcomes construct has not yet been identified, some measures are likely to be suitable in the interim. Thus, comprehensive real-life functioning assessment, using self-report, informant report and interviewer best judgment across six different real-life functioning rating scales may be required to capture the complexity of functional outcome in schizophrenia 58. The Specific Levels of Functioning (SLOF) Scale 62 is a 43-item multidimensional behavioral survey administered in person to the caseworker or caregiver, selected on the basis of his/her familiarity with that person or a patient-administered scale completed with verbal instructions from the examiner to rate its own performance. The scale does not include items relevant to psychiatric symptomatology or cognitive dysfunctions, but assesses the patient’s current functioning and observable behavior, as opposed to inferred mental or emotional states, and focuses on a person’s skills, assets, and abilities rather than deficits that once served as the central paradigm guiding assessment and intervention for persons with disabilities. It comprises six subscales: (1) physical functioning, (2) personal care skills, (3) interpersonal relationships, (4) social acceptability, (5) activities of community living and (6) work skills. The work skills domain comprises behaviors important for vocational performance, but is not a rating of behavior during employment. The latter would not be feasible, since the majority of patients with schizophrenia are unemployed; therefore, the proxy measure of work skills from the SLOF is used. Lastly, the SLOF also includes an open-ended question asking the informant if there are any other areas of functioning not covered...
by the instrument that may be important in assessing functioning in this patient. Each of the questions in the above domains is rated on a 5-point Likert scale. Scores on the instrument range from 43 to 215. The higher the total score, the better the overall functioning of the patient. According to the original version of the SLOF, the time frame covered by the survey is the past week. Each informant is asked to rank how well they know the patient on a 5-point Likert scale ranging from “not well at all” to “very well.” Ratings on individual items of the SLOF may be used to capture the current state of overall functioning while showing specific areas of therapeutic and rehabilitative need, i.e. to identify goals in planning treatment for clients, to develop special intervention or skill-training programs, or to assign clients with similar or complementary strengths and needs to existing programs. An adaptation of the SLOF is to allow patients to rate themselves on each item, while staff make independent judgments. Patients and staff then share their ratings, discuss discrepancies and negotiate a mutually acceptable set of functionally oriented goals for the plan. This process also could serve as a form of quality assurance, allowing patients and staff to obtain potentially valuable feedback about the patients’ self-perceptions and help staff to gauge better the accuracy of their judgments. Lastly, the SLOF has direct applications in research on patient outcome and program evaluation. The SLOF was found to be a reliable and valid scale, with a good construct validity and internal consistency, as well as a stable factor structure. In the context of the NIRP, the instrument was translated in Italian and its construct validity, internal consistency and factor structure as explored.

**Conflict of Interest**

None.

**References**


Sabbag S, Twamley EM, Lea Vella MA. Assessing everyday functioning in schizophrenia: not all informants see equally informative. Schizophr Res 2011;131:250-5.


