

Complete Methods

The data in this article are part of the five-site MATRICS Psychometric and Standardization Study (MATRICS PASS). Phase 1 was conducted with 176 schizophrenia patients tested twice at a 4-week interval (1), and phase 2 included 300 community subjects to collect co-norming data for the tests in the final battery (2). In addition to selecting the tests for the MATRICS Consensus Cognitive Battery, a second goal for phase 1 was to collect data on potential co-primary measures, which are the basis for the current article.

Human subject procedures were approved by each site's institutional review board. The subjects provided written informed consent after the procedures were fully explained. If the subjects were eligible for inclusion following diagnostic interview and case review, they were scheduled for baseline assessments. The subjects returned 4 weeks after the baseline assessment for a retest.

Assessments

The MATRICS PASS included several types of assessments, including ratings of clinical symptoms, cognitive performance measures, self-reports of community functioning, and potential co-primary measures. To evaluate the co-primary measures, we used a similar approach to that used to evaluate the cognitive performance measures. Specifically, we considered 1) test-retest reliability, 2) utility as a repeated measure, 3) relationship to functional status, 4) tolerability/practicality, and 5) number of missing data. In addition, an important

consideration for evaluating co-primary measures is the degree to which they correlate with cognitive performance (3). To evaluate the relationship with cognitive performance, we considered scores on the 10 tests that comprised the final battery (the MATRICS Consensus Cognitive Battery).

A. Co-Primary Measures

Based on discussions at consensus meetings and recommendations of the MATRICS Outcomes Committee (A. Bellack, chair), two approaches to co-primary measures were considered: measures of functional capacity and self-report measures of cognition. Two representative measures were included from each approach. The Outcomes Committee acknowledged that there are rather few candidate measures and that measure development in this area is still relatively immature compared with neurocognition. Unlike the large consensus and data collection process that was used to select cognitive performance tests for the MATRICS Consensus Cognitive Battery, the selection of potential co-primary measures was based on expert recommendations from the committee.

1. The Maryland Assessment of Social Competence (4): This measure is a functional capacity measure of participants' ability to solve common problems in an interpersonal context (e.g., interacting with a health care worker). It consists of four 3-minute conversations between the subject and a research staff member who portrays (role plays) a person in the community. The staff member's responses are scripted and designed to put primary responsibility for maintaining the conversation on the subject. A 90-second practice scene is administered first to acclimate the subject to the procedure. Role-play scenarios consist of one involving initiating conversation with a casual acquaintance, two involving negotiation and compromise (e.g., asking for a second chance on a job), and one involving standing up for one's rights (e.g., talking to a landlord about a leaky roof). The interactions are videotaped for later scoring. The procedure takes about 20 minutes to administer. Each scenario was coded on three dimensions on a 5-point Likert scale: verbal skill

(a measure of speech content), nonverbal skill (a measure of paralinguistic style such as eye contact and gestures), and overall effectiveness (a composite measure of the ability to maintain focus and achieve the goal of the scenario).

2. The UCSD (University of California at San Diego) Performance-Based Skills Assessment (5): This measure 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 events that the person may encounter in the community. Administration of the assessment requires an average of 30 minutes to complete. Interrater reliability of ratings are excellent (intraclass correlation coefficient=0.91, $p < 0.001$). In addition to subscale scores, the UCSD Performance-Based Skills Assessment yields a summary score ranging from 0 to 100. For this study, the UCSD Performance-Based Skills Assessment was modified to make the test relevant to all regions of the United States; bills and other testing materials that were specific to the San Diego area were modified to be generic. One additional component of the UCSD Performance-Based Skills Assessment was included in this study: a medication management task that assessed the participant's ability to take four different medications according to prescribed instructions over the course of a day. This component was a secondary measure and was analyzed separately.
3. The Schizophrenia Cognition Rating Scale (6): This measure is an 18-item 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, motor skills) that were chosen because of their severity of impairment in many patients with schizophrenia and the demonstrated relationship of these cognitive deficits to impairments in aspects of functional outcome. Two examples of items from the Schizophrenia Cognition Rating Scale 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. The anchor points for each item focus on the degree of impairment and the degree to which the deficit impairs day-to-day functioning. Complete administration of the Schizophrenia Cognition Rating Scale included two separate sources of information that generated three different ratings: an interview with the patient, an interview with an informant of the patient (family member, friend, social worker, etc.), and a rating by the interviewer. The interviewer's rating reflected a combination of the two interviews incorporating the interviewer's observations of the patient. A global score based on the sum of the 18-items was calculated for each type of rating (patient, informant, and interviewer).

4. The Clinical Global Impression of Cognition in Schizophrenia (7): This measure is similar in format and structure to the Schizophrenia Cognition Rating Scale, and it also involves interviews with the patient and informant. In both measures, the interviewer provides an integrated rating based on both sources of information. The Clinical Global Impression of Cognition in Schizophrenia has a general background section and includes four major categories for evaluation: activities of daily living, neurocognitive state—category severity, global severity of cognitive impairment, and Global Assessment of Functioning. For the purposes of this study, we did not administer the section on activities of daily living because these ratings were captured in other measures. One key difference between the Clinical Global Impression of Cognition in Schizophrenia and the Schizophrenia Cognition Rating Scale is that the former obtains ratings of cognition at the level of the specific cognitive domain (i.e., the seven cognitive domains included in the MATRICS Consensus Cognitive Battery) as well as overall cognition; the Schizophrenia Cognition Rating Scale obtains ratings only for overall cognition. The Clinical Global Impression of Cognition in Schizophrenia uses a 7-point Likert scale for its ratings, with higher ratings indicating more impairment. To make the statistical comparisons of the two measures more comparable, we only used the global score (global severity) from the Clinical Global Impression of Cognition in Schizophrenia, although potentially the domain ratings are an additional useful feature. The Clinical Global Impression of Cognition in Schizophrenia also includes an additional global rating of cognitive function provided by the interviewer called the Global Assessment of Functioning—cognition in schizophrenia that is rated on a 100-point scale. This rating supplements the Clinical Global Impression of Cognition in Schizophrenia global severity rating, and it parallels the DSM-IV Global Assessment of Functioning scale but with an emphasis on the impact of cognition on everyday functioning level.

The two functional capacity measures and the two interview-based measures were counterbalanced between, but not within, subjects. That is, each participant received the same order at baseline and retest.

Training and Quality Assurance for Co-Primary Measures

Training for the functional capacity measures (the Maryland Assessment of Social Competence and the UCSD Performance-Based Skills Assessment) included on-site training by the developers of the instruments (the University of Maryland for the Maryland Assessment of Social Competence, and the University

of California, San Diego, for the UCSD Performance-Based Skills Assessment). The training took 1.5 days total for both measures. Raters for the Maryland Assessment of Social Competence were selected from staff members who had experience administering standardized psychological tests. They were trained using role-played exercises with the developers, as well as tapes from a tape library. Raters on the UCSD Performance-Based Skills Assessment practiced administration with the testing materials under simulated conditions. Training included group exercises as well as one-on-one certification. Scoring of the UCSD Performance-Based Skills Assessment is objective and occurred at the time of testing, so it did not involve subsequent quality assurance checks. For the Maryland Assessment for Social Competence, each site sent tapes of two administrations with practice subjects to the University of Maryland for a final check on administration procedures. Afterward, the tapes of each administration of the Maryland Assessment of Social Competence were sent to the University of California at Los Angeles (UCLA) for scoring. Interrater reliability for each behavioral code of the Maryland Assessment of Social Competence was assessed by having a second rater conduct independent ratings of 20% of the group. The reliability (intraclass correlation coefficient) for the effectiveness rating across the four scenes was 0.69.

For the interview-based assessments (the Schizophrenia Cognition Rating Scale and the Clinical Global Impression of Cognition in Schizophrenia), initial training was provided on both scales in a one-day training session. Interviewers for these measures were selected from staff members who had

experience with semistructured psychiatric interviews or symptom rating scales. The training was conducted by the developers of the scales. Training on both measures used a combination of didactic materials and training videos that had consensus ratings. Following the in-person training, the quality assurance sites (Duke University for the Schizophrenia Cognition Rating Scale and UCLA for the Clinical Global Impression of Cognition in Schizophrenia) sent out two additional tapes with consensus ratings that were rated by the interviewers at each site to check for consistency in ratings. Finally, each site provided tapes of each type of interview to the quality assurance sites. These tapes were reviewed by the developers for ratings and interview style, and feedback was provided to each of the sites.

B. MATRICS Consensus Cognitive Battery

The MATRICS Consensus Cognitive Battery includes 10 tests from seven different cognitive domains: 1) the Trail Making Test Part A, 2) the Brief Assessment of Cognition in Schizophrenia: symbol coding, 3) the Hopkins Verbal Learning Test—Revised, and 4) the Wechsler Memory Scale—III: spatial span, 5) the Letter-Number Span, 6) the Neuropsychological Assessment Battery: mazes, 7) the Brief Visuospatial Memory Test—Revised, 8) Category Fluency (animal naming), 9) the Mayer-Salovey-Caruso Emotional Intelligence Test: managing emotions, and 10) the Continuous Performance Test—identical pairs.

C. Community Functioning

As described more fully in the first article (4), variables from the Birchwood Social Functioning Scale (8), supplemented with work and school items from the Social Adjustment Scale (9), were reduced through a principal-components analysis into three domain scores (factor scores for work, social, and independent living) as well as a summary score of global community functioning.

Blinding

Assessments at each site were conducted by three types of individuals. Testers at each site administered the measures of cognitive performance and functional capacity. Both types of tests are relatively objective in administration and scoring and designed to be administered by testers familiar with structured psychological tests. Interviewers (staff members different from the testers) conducted the two measures of self-reports of cognitive function. Finally, the site coordinators conducted the interviews for community status. Hence, there was a clear separation among the individuals who assessed functional capacity/cognitive performance, functional status, and self-reports of cognition. However, self-report interviews of cognition involve questions about the degree to which cognitive impairment affects daily functioning. Hence, the interview-based assessments of cognition could not be fully blind to functional status.

References

1. Nuechterlein KH, Green MF, Kern RS, Baade LE, Barch D, Cohen J, Essock S, Fenton WS, Frese FJ, Gold JM, Goldberg T, Heaton R, Keefe RSE, Kraemer H, Mesholam-Gately R, Seidman LJ, Stover E,

- Weinberger D, Young AS, Zalcman S, Marder SR: The MATRICS Consensus Cognitive Battery: part 1, test selection, reliability, and validity. *Am J Psychiatry* (in press)
2. Kern RS, Nuechterlein KH, Green MF, Baade LE, Fenton WS, Gold JM, Keefe RSE, Mesholam-Gately R, Mintz J, Seidman LJ, Stover E, Marder SR: The MATRICS Consensus Cognitive Battery: part 2, co-norming and standardization. *Am J Psychiatry* (in press)
 3. Buchanan RW, Davis M, Goff D, Green MF, Keefe RSE, Leon AC, Nuechterlein KH, Laughren T, Levin R, Stover E, Fenton W, Marder SR: A summary of the FDA-NIMH-MATRICES workshop on clinical trial design for neurocognitive drugs for schizophrenia. *Schizophr Bull* 2005; 31:5–19
 4. Bellack AS, Sayers M, Mueser K, Bennett M: Evaluation of social problem solving in schizophrenia. *J Abnorm Psychol* 1994; 103:371–378
 5. Patterson TL, Goldman S, McKibbin CL, Hughs T, Jeste DV: UCSD Performance-Based Skills Assessment: development of a new measure of everyday functioning for severely mentally ill adults. *Schizophr Bull* 2001; 27:235–245
 6. Keefe RS, Poe M, Walker TM, Kang JW, Harvey PD: The Schizophrenia Cognition Rating Scale: an interview-based assessment and its relationship to cognition, real-world functioning, and functional capacity. *Am J Psychiatry* 2006; 163:426–432

7. Ventura J, Cienfuegos A, Boxer O, Bilder RM: Clinical Global Impression of Cognition in Schizophrenia (CGI-CogS): reliability and validity of a co-primary measure of cognition. *Schizophr Res* (in press)
8. Birchwood M, Smith J, Cochran R, Wetton S, Copestake S: The Social Functioning Scale: the development and validation of a new scale of social adjustment for use in family intervention programs with schizophrenic patients. *Br J Psychiatry* 1990; 157:853–859
9. Weissman M, Paykel E: *The Depressed Woman: A Study of Social Relationships*. Chicago, University of Chicago Press, 1974

Table 1. Test-Retest Reliability of Measures in the MATRICS Psychometric and Standardization Study

Co-Primary Measures	Test-Retest Reliability (r)	Test-Retest Reliability (intraclass correlation coefficient)
Maryland Assessment of Social Competence		
Conversational content	0.68	0.68
Nonverbal content	0.66	0.66
Effectiveness	0.69	0.69
UCSD Performance-Based Skills Assessment		
Total score	0.70	0.70
Medication management	0.49	0.48
Schizophrenia Cognition Rating Scale		
Overall impression (patient)	0.60	0.60
Overall impression (informant)	0.77	0.77
Overall Impression (interviewer)	0.82	0.82
Clinical Global Impression of Cognition in Schizophrenia		
Neurocognitive state (patient)	0.76	0.76
Neurocognitive State (caregiver)	0.85	0.85
Neurocognitive State (composite)	0.80	0.80
Global Assessment of Functioning (composite)	0.89	0.89

Table 2. Relationship to Cognitive Performance in the MATRICS Psychometric and Standardization Study^a

Co-Primary Measures	Trail Making Test Part A	Brief Assessment of Cognition in Schizophrenia—Symbol Coding	Hopkins Verbal Learning Test—Revised	Wechsler Memory Scale—III—Spatial Span	Letter-Number Span	Neuropsychological Assessment Battery—Mazes	Brief Visuospatial Memory Test—Revised	Category Fluency (animal naming)	Mayer-Salovey-Caruso Emotional Intelligence Test—Managing Emotion	Continuous Performance Test—Identical Pairs	Overall Cognition
Maryland Assessment of Social Competence											
Conversational content	-0.12	0.29	0.40	0.16	0.37	0.08	0.17	0.26	0.34	0.33	0.42
Nonverbal content	-0.02	0.16	0.33	0.14	0.26	-0.03	0.12	0.23	0.27	0.24	0.32
Effectiveness	-0.10	0.26	0.38	0.14	0.32	0.07	0.16	0.26	0.34	0.30	0.40
UCSD Performance-Based Skills Assessment											
Total score	-0.36	0.46	0.50	0.36	0.48	0.36	0.39	0.41	0.31	0.51	0.61
Medication ^b	0.19	-0.41	-0.41	-0.44	-0.44	-0.29	-0.35	-0.33	-0.28	-0.33	-0.53
Schizophrenia Cognition Rating Scale											
Patient ^b	0.05	-0.14	-0.03	0.02	-0.12	-0.17	-0.14	-0.02	-0.08	-0.23	-0.18

Informant ^b	0.18	-0.20	-0.11	0.01	-0.25	-0.12	-0.22	-0.12	-0.16	-0.16	-0.21
Interviewer ^b	0.27	-0.33	-0.17	-0.05	-0.30	-0.24	-0.23	-0.14	-0.23	-0.27	-0.31
Clinical Global Impression of Cognition in Schizophrenia											
Neurocognitive state (patient) ^b	0.23	-0.28	-0.15	-0.02	-0.25	-0.30	-0.17	-0.15	-0.22	-0.30	-0.30
Neurocognitive state (caregiver) ^b	0.23	-0.27	-0.14	0.04	-0.24	-0.09	-0.14	-0.13	-0.20	-0.31	-0.24
Neurocognitive state (composite) ^b	0.30	-0.28	-0.16	-0.04	-0.28	-0.24	-0.19	-0.18	-0.24	-0.32	-0.31
Global Assessment of Functioning (composite)	-0.28	0.33	0.14	0.01	0.22	0.25	0.26	0.16	0.18	0.28	0.30

^aCognitive performance scores that were used were the following: Trail Making Test Part A: time to completion; Brief Assessment of Cognition in Schizophrenia symbol coding: total number correct; Hopkins Verbal Learning Test—Revised: total number of words recalled over three learning trials; Wechsler Memory Scale—III: spatial span: sum of raw scores on forward and backward conditions; Letter-Number Span: number of correct trials; Neuropsychological Assessment Battery: mazes: total raw score; Brief Visuospatial Memory Test—Revised: total recall score over three learning trials; Category Fluency (animal naming): total number of animals named in 60 seconds; Mayer-Salovey-Caruso Emotional Intelligence Test: managing emotions: branch score using general consensus scoring; Continuous Performance Test—Identical Pairs: mean d value across 2-, 3-, and 4-digit conditions.

^bTests in which lower scores are better.