Neurocognitive Assessment in the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) Project Schizophrenia Trial: Development, Methodology, and Rationale

by Richard S.E. Keefe, Richard C. Mohs, Robert M. Bilder, Philip D. Harvey, Michael F. Green, Herbert Y. Meltzer, James M. Gold, and Mary Sano

Abstract

Patients with schizophrenia are severely impaired in crucial aspects of neurocognitive function. This impairment is the strongest clinical correlate of poor long-term outcome and adaptive dysfunction. Reports of neurocognitive enhancement with second generation antipsychotic medications have thus offered promise for improvement in the long-term outcome of patients with schizophrenia. However, the majority of these studies have had serious weaknesses in methodology, such as open-label design, small samples, or inappropriate dosing of medications. More recent studies have addressed these methodological issues but have been of short duration and have largely been sponsored by pharmaceutical companies. The Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) project is a unique opportunity to address the comparative neurocognitive effectiveness of available antipsychotic medications. This article describes the neurocognitive methods used in the schizophrenia trial of the CATIE project, including the selection and training of neurocognitive raters, patient inclusion criteria for assessment, rationale for the choice of neurocognitive instruments, and methodology for each neurocognitive test.

Keywords: Neurocognitive function, cognition, antipsychotic medications, neuropsychology, schizophrenia, clinical trials.


Neurocognitive functioning in schizophrenia is severely impaired. In several landmark studies, the severity of cognitive impairment has reached two standard deviations below the mean in crucial aspects of cognitive functions (e.g., Saykin et al. 1991). Even meta-analytic techniques, which tend to underestimate the size of a statistical effect, have suggested that patients with schizophrenia perform more than a full standard deviation below the normal mean in numerous areas of neurocognitive ability (Heinrichs and Zakanis 1998). Neurocognitive impairment is present at the first episode of psychosis, and many cognitive domains may be as severely impaired in first episode patients as in patients with chronic schizophrenia (Hoff et al. 1992; Saykin et al. 1994; Bilder et al. 2000).

Neurocognitive impairment is associated with various aspects of schizophrenia. Patients with more severe cognitive deficits tend to have more severe negative symptoms (Manschreck et al. 1985; Addington et al. 1991; Strauss 1993) and symptoms of disorganization (Spitzer 1993), and worse adaptive dysfunction (Green 1996). While the correlations between severity of general cognitive deficits and positive symptoms have usually been nonsignificant or negative, some aspects of cognitive impairment, such as working memory, may also be associated with the severity of positive symptoms (Bressi et al. 1996; Carter et al. 1996). Most important, the severity of cognitive deficits in schizophrenia is associated with various aspects of poor outcome, such as the inability to acquire skills, poor social problem-solving, and poor community functioning (Green 1996; Green et al. 2000). In fact, cognitive impairment may be a better predictor of poor outcome than any other symptom domain (Green 1996).

Nearly 50 years of research has indicated that first generation antipsychotics at standard doses provide little benefit with regard to the cognitive disturbances of patients with schizophrenia (Medalia et al. 1988; Cassens et al. 1990; Blyler and Gold 2000). However, second gen-
eration antipsychotic medications may significantly improve cognitive performance in patients with schizophrenia. Clozapine (Buchanan et al. 1994; Lee et al. 1994), risperidone (Green et al. 1997; Kern et al. 1998; Purdon et al. 2000), and olanzapine (Meltzer and McGurk 1999; Purdon et al. 2000) treatments are associated with improved cognitive performance in patients with schizophrenia. A meta-analysis of studies assessing the cognitive effects of second generation antipsychotics, primarily clozapine or risperidone, suggested significant improvement in cognition with the second generation medications even when the results of each study were corrected for multiple comparisons (Keefe et al. 1999). Data published after this meta-analysis was completed suggest that olanzapine and quetiapine may have cognitive performance-enhancing properties that are at least as substantial as those reported with risperidone and clozapine (Sax et al. 1998; Meltzer and McGurk 1999; Velligan and Miller 1999; Purdon et al. 2000; Purdon et al., in press; Keefe et al., submitted). It should be noted, however, that the effect sizes for neurocognitive improvement with second generation antipsychotics compared to first generation antipsychotics are estimated to be small to medium (Harvey and Keefe 2001), and some have questioned whether these relative improvements with the second generation drugs are not secondary to other clinical changes, such as decreased extrapyramidal symptoms (EPS), decreased anticholinergic use, and improved negative symptoms that are associated with freedom from first generation drugs (Harvey and Keefe 2001; Carpenter and Gold 2002).

One of the goals of the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) project is to determine the relative impact of the available antipsychotic medications on relevant outcomes in schizophrenia, particularly with regard to those outcomes that result in increased effectiveness. Because cognitive impairment has been found to be strongly associated with functional ability, it is possible that cognitive improvement may be one of the key elements of long-term effectiveness. Surprisingly, few studies have demonstrated a relationship between treatment-related cognitive improvement and long-term improvement in functional outcome or quality of life (Buchanan et al. 1994; Meltzer and McGurk 1999), and no study with double-blind methodology or random assignment of medication has demonstrated that any of the available antipsychotic medications improve cognitive function in patients in general, with those improvements significantly correlated with better long-term outcomes.

The type of study that would be required to assess these aspects of effectiveness longitudinally with a sample size substantial enough to generate adequate power is very rare. The CATIE project presents an ideal opportunity to determine the relative impact of antipsychotic medications on cognition and to evaluate the impact of cognitive improvement (or worsening) on measures of treatment effectiveness.

In this article, we describe the rationale, development, and methodology for assessing neurocognitive function in the CATIE project.

**Neurocognitive Advisory Group**

The choice of tests for a neurocognitive battery is often controversial. Because this project promises to yield a potentially rich dataset, it was important that the batteries of tests chosen for the project be approved by leaders in schizophrenia and dementia research. It was determined that approval of the batteries and the neurocognitive approach for the trials would be particularly important given that numerous independent investigators may desire access to the neurocognitive data base for ancillary studies. Thus, the first step in determining the methodology for neurocognitive assessment in the CATIE project was to appoint and convene a Neurocognitive Advisory Group (NAG), in concert with National Institute of Mental Health (NIMH) staff and the External Scientific Advisory Board. NAG comprises the following individuals: Richard Keefe (director), Richard Mohs (formerly codirector, currently a consultant to NAG), Robert Bilder, Terry Goldberg, Michael Green, Philip Harvey, Herbert Meltzer, and Mary Sano. James Gold also joined the meeting as an ad hoc member to help align the neurocognitive battery of the CATIE project with the battery of tests used for the NIMH Treatment of Negative Symptoms and Cognitive Impairments project, which is investigating the cognitive effects of glycine and d-cycloserine in patients with schizophrenia. This group met on January 10, 2000, to discuss the research design and the batteries of tests to be used for the schizophrenia and dementia clinical trials. In addition, because various methodological challenges are intrinsic to the neurocognitive component of clinical trial research, the consultants used this opportunity to discuss several relevant assessment issues: tester training and certification methods, the development of specific exclusion criteria for the neurocognitive component of the study, bilingual testing, optimal test batteries, and the overlap of test constructs between separate trials of different disorders.

**Assessment Training**

**Overview.** The aim of the CATIE Neurocognitive Assessment Unit training program is for testers and investigators to achieve a thorough understanding of the rationale and methods of the neurocognitive assessment protocol. Considerable preparation preceded the production of
manuals and training materials. The basic manuals and data forms not only document planned procedures for each study but also serve as source documents for training in-house and site staff. The proper training of site personnel responsible for collecting, editing, and transmitting data is essential for ensuring a high-quality study. Accordingly, prior to enrollment of subjects at any of the sites, the Neurocognitive Assessment Unit trained site investigators and staff on all aspects of study design and implementation of neurocognitive assessments. Our philosophy was that training in multisite studies is most effective when it occurs in close temporal proximity to those critical points in the study timeline when site staff must acquire specific knowledge and skills in order to implement study tasks. Based on protocol design and site participation in multiple protocols, approximately 100 sites required neurocognitive tester certification: 65 in the schizophrenia trial, and 35 in the dementia trial. To increase cost efficiency and to ensure timely startup, testers were also certified based upon their prior involvement in studies using similar neurocognitive instruments. This reduced individual tester evaluations to 85 testers.

Specifics of Training Plan. All potential neurocognitive testers completed a Tester Evaluation Form to ensure that they had sufficient experience to collect accurate cognitive data from the relevant patient populations. Potential testers were scheduled for group and individual training by the Neurocognitive Assessment Unit staff at the Investigator Meeting. To help approved testers prepare for the meeting and training, each site was sent copies of the neurocognitive assessment manual and instructional videotape prior to the meeting. The manual includes descriptions of how to administer and score each test, score sheets if required, and detailed instructions for preparing and sending the raw test materials for data entry. The videotape demonstrating the test procedures for each test in the battery serves as a training and reference tool. It has also been useful as a training aid for new testers who did not attend the initial training meetings. Testers were informed that prior to the collection of neurocognitive data at their site, they would need to be certified for their ability to complete each of the tests of the test battery.

Training Meeting. The details of the final batteries were presented during the Neurocognitive Assessment Training at the Investigator Meeting. Testers from each investigative site attended this meeting to learn, practice, and become expert in the administration and scoring procedures for the diagnosis-specific neurocognitive batteries. The meeting included a presentation of all testing and scoring procedures to the group. Following this presenta-

tion, testers had opportunities to practice the tests in small groups supervised by psychologists fully familiar with the entire battery. Although many of the testers had extensive experience with the tests in the battery, or with psychological assessment in general, all were required to pass certification procedures. The rationale of this certification was explained to testers: even highly qualified psychologists can develop idiosyncratic data collection methods that may thus differ from site to site. Testers were instructed that although their method may indeed be superior to the method established for this trial, consistency across sites is the overriding goal.

Certification of Testers. Testers were told that they would be expected to master the testing procedures prior to the initiation of the trial at their site. To verify mastery of all testing procedures, all testers met individually with a Neurocognitive Assessment Unit evaluator to determine whether they were satisfactorily proficient with the neurocognitive testing batteries. Each potential tester administered the complete neurocognitive battery to the neurocognitive evaluator, who demonstrated some of the challenges that could be expected when testing patients with schizophrenia. A list of particularly important challenges was made prior to the meeting as a reference for evaluators, so that each potential tester would be given the opportunity to be tested similarly. For example, testers were challenged by "patients" who deviated from the rules in various ways, or who had sudden decrements in performance. Testers who failed the initial certification procedures were given an opportunity to practice test administration and scoring procedures during the meeting, and then later attempt to be recertified. Certified testers were sent a document indicating that they were permitted to conduct neurocognitive testing for the CATIE project.

Following initial certification, each tester was required to send the computer test files, source documents, and summary score sheets from the first five completed assessments. Data collection and scoring procedures were audited by following each measure from the raw testing material or response sheet, through the relevant scoring procedures, to the recording of the data on summary score sheets, to the online data from the computer data base. These results were reviewed by the Neurocognitive Assessment Unit personnel, and any errors were corrected with a full explanation of the necessary changes. Testers were required to complete five consecutive testing sessions without error.

Postmeeting Certification. A small number of testers were not able to attend the initial Investigator Meeting. Arrangements were made for these testers to follow the same procedures as those who attended the meeting, with
the exception that Neurocognitive Assessment Unit personnel visited these sites to conduct the same certification procedures as had been completed at the meeting. These procedures and the necessary documents were made available to new sites via the CATIE website.

Several sites requested the certification of an additional tester to perform testing in the event of vacation, sickness, or other potential absence of the primary tester. All additional and replacement testers underwent the same evaluation procedures. These procedures, made available to all principal investigators and testers on the CATIE website, are as follows: complete a form assessing testing experience and send it to the Neurocognitive Assessment Unit; obtain a copy of the most recent version of the manual and review the manual, testing procedures, instructional video, and tests with the help of a Certified Tester; following intensive practice, test the Certified Tester; test an individual under the supervision of the Certified Tester; and finally, contact the Neurocognitive Assessment Unit to arrange a Phone-Questionnaire Interview.

**Initial and Ongoing Site Interaction and Assessment.** The performance of neurocognitive testers has been monitored during the study, and any administration and scoring problems have largely been addressed and resolved as they occur. Neurocognitive Assessment Unit staff are available on a daily basis by telephone or e-mail to answer questions or issues that may occur at the sites.

A few weaknesses in the clarity of the neurocognitive procedure instructions have been communicated to the CATIE Neurocognitive Assessment Unit throughout the course of the study. These issues have precipitated communications from the Unit to all testers and principal investigators in the form of letters or mass e-mails. In addition, the certification procedures and relevant updates are made available to testers on the CATIE website.

Raw data from each site are entered directly into the data base using a web-based data entry system. As an additional check on the quality of neurocognitive data, the Neurocognitive Assessment Unit has made random requests for testers to send the source documents and summary scores for patients tested in the study. These scores are compared to those recorded in the electronic data base for the study.

**Neurocognitive Assessment Procedures**

**Testers.** NAG recommended that two testers be used from each site and that the same tester should complete the assessments for a particular patient, especially at the baseline and 8-week assessments for the schizophrenia trial.

**Timing of Assessments.** The timing of neurocognitive assessments presented a unique challenge, because patients could potentially change phases of the study, and thus the antipsychotic medication they received, at least three times during their participation in the study (Stroup et al., this issue). The group weighed the merits of following an "effectiveness model" by maintaining a fixed assessment schedule based upon the initial randomization of a patient to a specific study arm, versus the merits of an "efficacy model" of renewing the time schedule each time a patient entered a new treatment phase of the study. The favorable component of the effectiveness approach was that the neurocognitive effects of the initial treatment randomization could be determined with parallel timing in all patients. The unfavorable component was that this approach would preclude direct determination of the neurocognitive effects of changes in antipsychotic medication that occurred in later phases of the study. However, the group also acknowledged that restarting the testing schedule each time a patient entered a new phase of the study was likely to produce a greatly variable baseline level of performance that could be explained by a patient's previous treatment and was thus not random. We decided on a mix of the two. Because little was (and is) known about the relative short-term impact of the available antipsychotics on cognition, the group determined that the effectiveness model would be followed, with additional assessments after 2 months of treatment in the second phase of the study. This strategy would enable a larger data base for investigating the relative impact of antipsychotics during a time period consistent with standard treatment choices.

While the baseline neurocognitive assessments will be affected by numerous factors such as substance abuse, previous polypharmacy, and variable clinical state, the group felt that baseline neurocognitive assessments should be collected on all patients when possible. Without baseline data, we would not be able to determine if differences across treatment groups at this stage were due to differential short-term effects on cognition, or due to baseline neurocognitive differences across groups. A relatively early followup assessment of 2 months would meet these two purposes.

In sum, the group suggested the following time sequence of neurocognitive assessments for the schizophrenia trial, based upon time from randomization: baseline, 2 months, 6 months, and 18 months or end of study. For patients who enter a second phase of the study, testing is also conducted after 2 months in the second phase. If there is temporal overlap of less than 1 month between an assessment that is based upon the initial schedule and the phase 2, month 2 assessment, only one of the assessments is performed.
Inclusion Criteria

Spanish-Speaking Subjects. The group decided that, when possible, it would be important to test patients who use Spanish as their primary language. A manual for Spanish-speaking subjects in the United States was developed with assistance from a cross-cultural neuropsychologist, Dr. Lidia Artiola. For those tests not available in Spanish, test instructions, materials, and procedures were translated into Spanish. Sites with testers who planned to test patients in Spanish were required to demonstrate testing competency in Spanish and English.

Assessment of Primary Language. In patients for whom English is not the primary language, the group determined that English testing could proceed if a sixth grade reading level using the Wide Range Achievement Test, third edition (WRAT-III), Reading test (Wilkinson 1993) was demonstrated.

Pervasive Developmental Disorder. The group determined that patients who have received a diagnosis of pervasive developmental disorder (mental retardation) should be excluded.

Adjunctive Medications. Regarding adjunctive medications during the trial, the group suggested that medications used continually, such as anticholinergics, were acceptable. Medications used infrequently, such as onetime use of benzodiazepines for anxiety, might interfere with test performance, and were not to be given on the day of testing.

Previous Testing Experience. Patients who have participated in neurocognitive research studies with repeated assessments of the measures in this study might perform at higher than expected levels. An assessment of previous experience (within the past 6 months) with the tests in the neurocognitive battery was included in the neurocognitive protocol.

Neurocognitive Assessment Battery

The proposed neurocognitive assessment battery for this trial was developed as a result of the meetings of NAG and the Neurocognitive Assessment Unit. A number of guidelines were used to facilitate the selection of tests. These guidelines are based on the suggestions of Mohs (1995), Davidson and Keefe (1995), and Keefe et al. (1999) regarding the use of neurocognitive measures to assess treatment response. The guidelines include (1) adequate psychometric properties of the procedure, including minimal ceiling and floor effects in the diagnostic group being studied; (2) reliability in the diagnostic group being studied; (3) known deficits compared to normal controls in the diagnostic group being studied; (4) minimal test-retest practice effects and/or availability of alternate forms; (5) ease of administration to reduce examiner error across sites; (6) brevity; (7) relevance to the clinical population being studied; (8) relevance to specific hypotheses about the effect of second generation antipsychotic treatment on cognition in the group being studied; (9) relation to important outcome variables, such as community functioning; (10) previous suggestions that second generation or first generation antipsychotic treatments have a possible positive or negative impact on the functions measured by the test; and (11) if possible, available normative data.

Computerized Assessments. One discussion point among the group members was the risk/benefit of computerized versus pencil-and-paper testing. Several group members expressed concerns about computerized testing given their previous experiences with computerized tests often resulting in more missing data points than noncomputerized tests. We ultimately chose a combination of the two methods. To minimize some of the difficulties caused by tests administered from variable computers and software platforms, we purchased 65 desktop computers, developed and installed the software for the computerized tests at the Neurocognitive Assessment Unit, and shipped the test-ready computers to each site. This method allowed us to maintain consistent computerized test conditions across sites and ensured that novel software-hardware interaction difficulties would be minimized throughout the course of the study.

Neurocognitive Assessments

The following assessments are completed at the baseline visit only: education level; previous experience with the tests in the neurocognitive battery; and the WRAT-III Reading subtest. To complete the WRAT-III Reading subtest, patients simply read aloud a series of letters and words of increasing difficulty. If for some reason these assessments cannot be completed at the baseline visit, they can be completed at a subsequent visit.

The following neurocognitive tests are given at each neurocognitive visit in the order listed. The battery is estimated to take an average of 90 minutes.

- Controlled Oral Word Association Test (Phonological Fluency)
- Category Instances (Category Fluency)
- Wechsler Intelligence Scale for Children—third edition (WISC-III) Mazes
- Letter-Number Span Test
Verbal Fluency. Verbal fluency is severely impaired in patients with psychotic disorders and dementia. One of the most robust effects of second generation antipsychotic medication on cognition is the improvement of verbal fluency (Keefe et al. 1999; Harvey and Keefe 2001). Even after the analyses from all studies were corrected for multiple comparisons, four of six studies demonstrated significant improvement on verbal fluency measures with clozapine.

Controlled Oral Word Association Test (Benton and Hamshcer 1978). Subjects are asked to generate as many words as possible that start with a given letter (F, A, or S) in each of three 60-second trials. Measures: number of correct words generated in each category.

Category Instances (Benton and Hamshcer 1978). In three separate trials, subjects are given 60 seconds to generate as many words as possible within the categories of animals, fruits, and vegetables. Measures: number of correct words generated in each category (Benton and Hamshcer 1978).

Working Memory. Working memory has been described as a fundamental aspect of cognition (Goldman-Rakic 1987). Recent developments in the understanding of prefrontal functions in humans have followed extensive work on the neural circuitry underlying working memory function and dysfunction. Although conventional antipsychotics block dopamine receptors in the prefrontal cortex, and this has been found to impair working memory functions in nonhuman primates under normal conditions (Sawaguchi and Goldman-Rakic 1994), conventional antipsychotics do not impair or improve working memory functions in patients with schizophrenia when given at usual doses (see reviews by Medalia et al. 1988; Cassens et al. 1990; Goldberg and Weinberger 1996). The evidence to support clozapine-related enhancement of working memory functions is weak. While one study reported improvement in all cognitive tests, including auditory consonant trigrams, with clozapine treatment (Galletly et al. 1997), other work suggests that treatment with clozapine does not improve verbal working memory as assessed by auditory consonant trigrams (Hagger et al. 1993; Lee et al. 1994) or digits backward (Grace et al. 1996). Risperidone may improve verbal, visuospatial, and central executive aspects of working memory functions. Risperidone treatment had a greater beneficial effect than haloperidol treatment on verbal working memory as assessed by a digit span distraction test (Green et al. 1997) and digits backward (Rossi et al. 1997). Data from Meltzer and McGurk (1999) suggest that in patients with schizophrenia, risperidone may also improve visuospatial working memory functions. Thus, preliminary work indicates that various aspects of working memory may be improved by treatment with risperidone. The neurocognitive battery includes two measures of working memory.

Letter-Number Span test of auditory working memory (Gold et al. 1997). Patients are presented aurally with clusters of letters combined with numbers (e.g., N6G2). They are asked to reorder the cluster and tell the experimenter the numbers first, from lowest to highest, then the letters in alphabetical order. Measure: number of correct sequences.

Computerized test of visuospatial working memory (Hershey et al. 1999). Subjects must focus on a central fixation cross on a computer screen. While the cross is fixated, a cue appears for 150 ms in one of 32 possible locations at a 4.5-inch radius from the central fixation. A delay period (5 or 15 seconds) is then imposed. During the delay, a series of geometric shapes appears in place of the fixation cross. The subject must press the spacebar whenever the diamond shape appears. After the delay, the fixation cue returns, and the subjects must point on the computer screen to where they remember seeing the cue. On the cue-present trials the cue is present during the response phase. This set of trials gives an indication of subjects' pointing accuracy. Mean error in millimeters (distance between recall and actual target) is calculated for each subject for each type of trial. There are eight trials at each delay and eight cue-present trials. Measures: mean error for each type of trial.

Verbal Learning and Memory. Verbal memory is severely impaired in schizophrenia (Saykin et al. 1991) and dementia (Mohs et al. 1997) and is significantly associated with outcome in patients with schizophrenia (Green 1996). Improvement in verbal memory with any of the second generation antipsychotic medications would be of great value. One study showed improvement of declarative verbal learning/memory with risperidone compared to clozapine and haloperidol (Bilder et al. 2002), while an advantage for olanzapine over haloperidol and risperidone was suggested in another study (Purdon et al. 2000). Verbal memory is
most often assessed with measures of recall of stories or lists of words. An important feature of memory testing in longitudinal treatment studies is the use of alternate forms to minimize confusion between practice effects and genuine treatment effects. Alternate forms are available for tests of verbal memory and learning of lists of words.

**Hopkins Verbal Learning Test (Brandt 1991).** This test consists of 12 nouns read aloud for three consecutive trials, each trial followed by a free-recall test. Six alternate forms are used and counterbalanced across sites and across assessment periods. Measures: number of items recalled on each trial.

**Social Cognition.** The cognitive processes involved in how people perceive, interpret, and process information about themselves, others, and social situations are commonly impaired in persons with schizophrenia (Penn et al. 1997), which may contribute to impairments in a variety of social outcomes (Corrigan and Toomey 1995; Penn et al. 1996; Mueser et al. 1997). Social cognition may be a key intervening variable between basic cognition and functional outcome (Green and Nuechterlein 1999) and may be an important intervention target. Despite the importance of this component of cognition, no data have assessed its response to pharmacological treatment.

**FEDT (Kerr and Neale 1993).** Facial emotion discrimination is assessed with the FEDT, which requires the subject to determine whether two faces presented next to one another are expressing the same or different emotions. Thirty pairs of target faces are presented to the patient. Measures: the number of faces correctly discriminated.

**Motor Function.** Motor functions have been found to improve with various second generation antipsychotics (Gallhoffer et al. 1996; Myer-Lindenberg et al. 1997; Purdon et al. 2000). In addition to the direct measures of motor function, the cognitive functions that are most responsive to second generation antipsychotics have a timed component. This pattern may be a result of the absence of EPS from second generation antipsychotic medications. That is, because timed tests all involve some degree of dependence upon motor skills, which are impaired by EPS, the advantage of second generation antipsychotics could partially be a result of the reduced EPS. Thus, it is important to include tests of motor functions in the battery. Furthermore, motor functions are related to outcome (Bilder et al. 1985), underscoring the importance of this domain. A Grooved Pegboard test is used to measure motor function, and the Digit Symbol Test of the WAIS–R is used to measure a related construct, graphomotor speed.

**Grooved Pegboard (Lafayette Instrument Company 1989).** Patients are asked to insert in a specified order 25 pegs with keys into a pegboard with randomly positioned slots. Two 45-second trials are completed with the dominant hand. Measure: average number of pegs successfully inserted.

**WAIS–R Digit Symbol Test (Wechsler 1974).** Each numeral (1 through 9) is associated with a different simple form. Patients are given a list of numerals and are asked to copy as many forms associated with the numerals as possible in 90 seconds. Measure: number of forms accurately copied.

**Attention.** Attention is a fundamental cognitive deficit in patients with schizophrenia (Nuechterlein and Dawson 1984) and is associated with outcome in these patients (Green 1996). It is one of the few measures that demonstrate some improvement with first generation antipsychotic medications (Medalia et al. 1988; Blyler and Gold 2000). In previous studies, attention has been found to improve with risperidone (Stip and Lussier 1996; Rossi et al. 1997; Kern et al. 1998; Bilder et al. 2002), clozapine (Zahn et al. 1994; Grace et al. 1996), and olanzapine (Meltzer and McGurk 1999; Purdon et al. 2000; Bilder et al. 2002; Keefe et al., submitted). Attention is traditionally measured with the CPT.

**Computerized CPT.** The Identical Pairs version of the CPT (Cornblatt et al. 1988) has high test-retest reliability, making it ideal for studies with repeated assessments. CATIE uses a version of the CPT that includes three 150-trial conditions of increasing difficulty. This procedure ensures that data can be collected on the more cognitively impaired patients, yet very few patients perform perfectly under the most difficult conditions. Each condition involves the presentation of stimuli on a computer screen at the rate of one per second. In the first condition, two-digit numbers are presented and the subject lifts his or her finger whenever the two-digit number is a repeat of the previous two-digit number. The second and third conditions are the same as the first condition, except that the numbers are three digits and four digits, respectively. This test is administered on a desktop computer with a high-resolution monitor and an external mouse.

**Executive Function.** Various components of executive functions are impaired in patients with schizophrenia (Goldberg et al. 1987; Saykin et al. 1991), including set shifting, abstraction, problem-solving, and allocation of cognitive resources (Bressi et al. 1996). Two of the most frequently cited components of executive function deficit in schizophrenia are card-sorting (Goldberg et al. 1987) and completion of mazes (Gallhofer et al. 1996).

**WCST, 64-card computerized version (WCST–64P) (Heaton et al. 1993).** Performance on the WCST has been found to improve with clozapine, risperidone, and olanza-
pine, although many negative findings have been reported (reviewed in Meltzer and McGurk 1999). This area of cognition is particularly challenging in clinical trials with repeated assessments because a usual component of a measure of executive function is to learn the rules of the test. Once these rules are learned, they are often remembered, even after a long (e.g., 6-month) delay period. This issue is highlighted through the use of the WCST. Subjects sort a series of stimulus cards by matching them to four “key cards” that differ by form, color, and number. Successful performance on the WCST depends upon learning how to sort the cards and how to switch the sorting strategy when appropriate, because the “correct” sorting strategy changes after 10 consecutive correct responses. There is some controversy as to whether patients with schizophrenia benefit from previous exposure to the WCST. An early study with very chronic subjects suggested that patients with schizophrenia did not learn from previous assessments with the WCST (Goldberg et al. 1987). However, data from patients with heterogeneous cognitive performance (Green et al. 1990; Bellack et al. 1996) suggest that the performance of many patients with schizophrenia does improve with repeated exposure to the test. This issue may be particularly important in this project, because the mechanism by which second generation antipsychotic medications improve cognition to a greater extent than first generation antipsychotics may be related to the improvement of episodic memory, which may help patients recall how they performed the test on previous testing sessions. Subcortical dopaminergic blockade may inhibit this improvement (Robbins 1990; Robbins et al. 1990). This notion has been supported by data suggesting that schizophrenia patients treated with risperidone continue to benefit from repeated exposure to a test of processing capacity, while patients treated with haloperidol are limited in the gains they receive from additional administrations of the test (Harvey et al. 2000). Thus, the sole use of a test such as the WCST to measure executive function in this study may lead to confusion between improvements in executive function and improvements in episodic memory. Furthermore, differences in performance resulting from the different second generation medications may be difficult to detect if a sizable percentage of patients learn the test well enough to perform at near-perfect levels. Thus, we chose a 64-card version of the WCST for this study to minimize learning effects and decided to administer it on a desktop computer to minimize scoring errors. A 64-card computerized version of the WCST had been used with success in a previously completed industry trial. Measures: number of perseverative errors, completed number of categories and additional consecutive cards in the final category.

**WISC-III mazes (Wechsler 1991).** This is another test of executive function that is administered for this trial. The performance of patients with schizophrenia on maze tasks, usually impaired in schizophrenia, has been found to be improved with risperidone (Gallhoffer et al. 1996).

In this test, patients use a pencil to attempt to draw through a series of nine mazes without entering blind alleys. Performance is timed. Measure: raw score.

### References


Buchanan, R.W.; Holstein, C; and Brier, A. The comparative efficacy and long-term effect of clozapine treatment...


Purdon, S.; Malla, A.; Labelle, A.; and Lit, W. Neurocognitive change in schizophrenia after 6 months of double blind treatment with quetiapine or haloperidol. *Journal of Psychiatry and Neuroscience*, in press.


**Acknowledgments**

This article was based on results from the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) project supported with Federal funds from the National Institute of Mental Health (NIMH) under contract NO1 MH90001. The aim of this project is to examine the comparative effectiveness of antipsychotic drugs in conditions for which their use is clinically indicated, including schizophrenia and Alzheimer's disease. The project was carried out by principal investigators from the University of North Carolina, Duke University, the University of Southern California, the University of Rochester, and Yale University in association with Quintiles, Inc., and the program staff of the Division of Interventions and Services Research of the NIMH and investigators from 84 sites in the United States. AstraZeneca Pharmaceuticals LP, Bristol-Myers Squibb, Forest Pharmaceuticals, Inc., Janssen Pharmaceutical Products, L.P., Eli Lilly and Company, Otsuka Pharmaceutical Co., Ltd., Pfizer Inc., and Zenith Goldline Pharmaceuticals, Inc., provided medications for the studies.

**The Authors**

Richard S.E. Keefe, Ph.D., is Director, Neurocognitive Assessment Unit, Clinical Antipsychotic Trials of Intervention Effectiveness project; and Associate Professor of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC. Richard C. Mohs, Ph.D., is Lilly Research Fellow, Eli Lilly and Company, Indianapolis, IN. Robert M. Bilder, Ph.D., is Chief of Medical Psychology-Neuropsychology, Neuropsychiatric Institute and Hospital, Geffen School of Medicine, University of California, Los Angeles. Philip D. Harvey, Ph.D., is Professor of Psychiatry, Mt. Sinai School of Medicine. Michael F. Green, Ph.D., is Professor in Residence, Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, CA. Herbert Y. Meltzer, M.D., is Bixler Professor of Psychiatry, Vanderbilt University School of Medicine, Nashville, TN. James M. Gold, Ph.D., is Associate Professor of Psychiatry, Maryland Psychiatric Research Center, University of Maryland School of Medicine, Baltimore, MD. Mary Sano, Ph.D., is Professor, Department of Psychiatry, Mount Sinai School of Medicine, Bronx Veterans Medical Research Center, Bronx, NY.