5

Computer-Based Cognitive Testing

Thomas H. Crook, Gary G. Kay, and Glenn J. Larrabee

More than a decade ago, in the second edition of this book, we described the general merits and limitations of using computers in cognitive testing and went on to describe a number of specific batteries available to clinicians and researchers (Larrabee & Crook, 1996). In looking back, we see that, like old soldiers, several of the test batteries we described have just faded away, while others have been further refined and strengthened, and yet others, many others, have emerged and may provide useful new tools in neuropsychological evaluation. Aside from these new tools, though, the decade has illustrated the limitations of computerized cognitive testing. Principally, that neuropsychological diagnostic testing without a neuropsychologist is probably not a great idea. A number of Web sites and test vendors have purported to do just that and have, thankfully, failed to find acceptance in the marketplace. That is not to say that computerized memory testing via the Internet is not appropriate for preliminary screening, and even repeated testing over time to detect change, but it is not now, and is not likely to become, a substitute for traditional clinical diagnostic evaluation. Moreover, any Internet-based testing should be under the supervision of a qualified psychologist to minimize the chances of misattribution of results and related iatrogenic factors (Mittenberg et al., 1992; Suhr & Gunstad, 2002).

There are so many new computerized cognitive test batteries that a complete review could occupy this entire volume. There is also a great deal of overlap between many batteries. Thus, we have chosen among the available test batteries those with the best psychometric foundations, and those that have taken a novel approach to testing. But, before the review, let us consider the value that computers can bring to clinical neuropsychological evaluation.

Advantages of Computers in Neuropsychology

Even two decades ago, computers were seen as having advantages in testing far beyond data storage. These were said to include data scoring and analysis, analysis of test profiles for diagnostic classification, increased reliability, enhanced capacity to generate complex stimuli, greater accuracy and superior time resolution, standardization of stimulus presentation, and ease of administration, which reduces the need for highly skilled personnel (Adams, 1986; Adams & Brown, 1986). Limitations were also recognized, and these were said to include alterations in the patient's perception of and response to the test when a standardized paper and pencil test is adapted for computerized administration (Adams & Brown, 1986) and differential familiarity of patients with computer manipulanda used in some computerized test batteries (Kapur, 1988). Larrabee and Crook (1991) also emphasized the need to thoroughly validate computerized batteries and not assume that a validated paper and pencil test remains valid when administered via computer, the need for alternate forms and extensive normative data, and the need to develop tests that are "ecologically valid," that is, relevant to the tasks patients perform in everyday life and the symptoms that underlie many neurocognitive disorders.

Although there are advantages to computerized cognitive assessment, there are also limitations. Adapting existing neuropsychological tests to computerized administration changes the nature of the tests. This, in turn, can influence the nature of the patient's perception of the test, which can affect his or her motivation and response style (Adams & Brown, 1986). Kapur (1988) cautions that most computerized measures have significant visuoperceptual demands, which can cause difficulty for patients with reduced visual acuity or neglect. In addition, many computerized test batteries require the patient to utilize manipulanda that may be unfamiliar, such as a mouse or computer keyboard. This may pose particular difficulty for severely impaired patients, especially among the elderly (Larrabee & Crook, 1991).

Finally, it is important that psychologists utilizing computers for evaluation adhere to professional standards regarding such instruments (Division 40 Task Force Report on Computer-Assisted Neuropsychological Evaluation, 1987; Matthews et al., 1991). Computerized neuropsychological evaluation, in general, and memory testing, in particular, should be conducted in line with APA guidelines for test instruments concerning reliability and validity, and user qualifications.

Two basic approaches have been employed in computerized cognitive assessment. The first, exemplified by the MicroCog (Powell et al., 1993), MindStreams (Dwolatzky et al., 2004), CogScreen (Kay, 1995), Cognitive Drug Research (CDR; Wesnes et al., 1987), Cambridge Neuropsychological Tests Automated Battery (CANTAB; Morris et al., 1987), ImPACT (Miller et al., 2007), Automated Neuropsychological Assessment Metrics (ANAM; Reeves et al., AQ1 2007), and Cog State (Maruff, 2007) batteries, adapts standard cognitive tasks for computerized administration and scoring. The second approach, exemplified by the Psychologix Battery developed by Crook, Larrabee, and Youngjohn (e.g., Larrabee & Crook, 1991), makes use of video and computer graphics technologies to simulate memory and other cognitive tasks of everyday life.

Computerized Batteries for Assessing Memory and Related Cognitive Abilities

MicroCog

MicroCog was developed in an effort to identify cognitive status changes in physicians and other professionals that might interfere with occupational performance (Kane, 1995; Powell et al., 1993), but it has subsequently become a general neuropsychological screening instrument. Since 2003, it has been available from Psychological Corporation on a Windows platform. The battery has not been widely accepted by either clinicians or researchers. Nevertheless, in recent years, the program was revised and now operates on a standard PC running the Windows operating system. Responses are entered using a standard keyboard, and this mode, of course, may pose a problem because some testees, particularly older males, may have limited experience with a keyboard.

Five primary neurobehavioral domains are assessed, including Attention/Mental Control, Memory, Reasoning/Calculation, Spatial Processing, and Reaction time. The Attention/Mental Control Domain includes span tasks (numbers forward and reversed), a continuous performance task (Alphabet subtest) requiring identification of letters of the alphabet as they appear in sequence within a series of random letters, and a supraspan word list test (Word Lists subtest) presented in continuous performance format, with subsequent assessment of incidental learning. Memory is assessed for immediate and delayed recognition of the content of two stories read by the examinee and delayed recognition of a street address, assessed in a multiple-choice (i.e., recognition) format. Reaction time assesses visual and auditory simple reaction time. The Spatial Processing domain also includes a visual working memory subtest, in which the subject must reproduce, following a one-second presentation, a grid pattern in a 3×3 matrix, in which three, four, or five of the spaces are colored.

MicroCog can be administered in the 18-subtest standard form in one hour or as a 12-subtest short form in 30 minutes. Scores are provided for accuracy, speed, and proficiency (a combination of speed and accuracy). Measures of general cognitive function and proficiency are also computed. There are only two alternate forms, and thus the battery is not appropriate for many research applications or, perhaps, longitudinal clinical follow-up.

A strength of MicroCog is the normative database of 810 adults (45 females and 45 males in each of nine age groups: 18–24, 25–34, 35–44, 45–54, 55–64, 65–69, 70–74, 75–79, 80–89). Test data are also presented for a variety of clinical groups, including dementia, lobectomy, depression, and schizophrenia. Data are provided for percentage correct classification, sensitivity, and specificity for each clinical group. In addition, mean levels of performance for the five neurobehavioral domain index scores, processing speed, processing accuracy, general cognitive function, and proficiency are provided for each clinical group.

Test–retest reliability does not appear to be a great strength of MicroCog, although it is difficult to know from the paper most frequently cited (Elwood, 2001) how it was established. It is not clear whether subjects were tested on different days, with different forms, and so on. In a proper test–retest study in a small sample of 40, normal subjects (Raymond et al., 2006) ranged from .49 to .84 at 2 weeks and from .59 to .83 at 3 months.

Factor analytic data are provided that demonstrate two factors: information-processing accuracy and information-processing speed (Powell et al., 1993). Concurrent validity data are provided that demonstrate the correlation of various MicroCog scales with external test criteria. For example, the MicroCog Attention/Mental Control Index correlates .57 with the Wechsler Memory Scale-Revised (WMS-R; Wechsler, 1987) Attention/Concentration Index and the MicroCog Memory Index correlates .44 with the WMS-R General Memory Index (Powell et al., 1993). These are not impressive correlations. Data relevant to construct validity are provided by Ledbetter and Hurley (1994). In a combined MicroCog WMS-R factor analysis, MicroCog Alphabet, Word List, and Numbers Forward and Reversed loaded on the same attention factor as WMS-R Digit Span and Visual Memory Span, whereas MicroCog immediate recognition/recall of two stories loaded on a memory factor with WMS-R Logical Memory I. A recent study (Helmes & Miller, 2006) among older subjects in the community found modest correlations between MicroCog and Wechsler Memory Scale-Three (WMS-III; Wechsler, 1997) subtests of the same construct. Correlations between the visual subtests of the two tests were not even statistically significant. So, it seems clear that MicroCog is not a substitute for the WMS-R.

Kane (1995) reviewed MicroCog and noted that its strengths included the computerization of a number of traditional neuropsychological measures, the addition of proficiency scores, the provision of detailed information on standard error of measurement for subtests and general performance indices, and sizable age- and education-based norms. Weaknesses were noted to be the use of a multiple key interface and lack of motor and divided attention tasks. Also, Kane and Kay (1992) noted that much of the psychometric data cited for MicroCog was actually obtained for earlier versions of the test.

MindStreams

MindStreams is an "Advanced Cognitive Health Assessment" battery developed by NeuroTrax, a company founded in Israel in 2000, specifically to develop and commercialize computerized cognitive testing. MindStreams was launched in 2003 and consists of a battery of mostly traditional neuropsychological tests administered and scored by computer. Tests are downloaded over the Internet and scores are transmitted back to a NeuroTrax central computer, where data are processed and scores calculated. Tests can be administered on a standard PC, and a combination mouse/key pad used by the subject in responding during testing is provided by MindStreams.

Individual test scores and index scores for Memory, Executive Function, Visual Spatial and Verbal Function, Attention, Information-Processing Speed, and Motor Skills are calculated and transmitted back to the user within minutes in the form of a detailed report. The report compares an individual's scores with scores in earlier test sessions and with those of subjects matched for age and education within a normative database of 1659 subjects between the ages of 9 and 95. Tests in the Global Assessment Battery include Verbal and Nonverbal Memory, Problem Solving (Nonverbal IQ), Stroop Interference, Finger Tapping, Catch Game, Staged Information-Processing Speed, Verbal Function, and Visual Spatial Processing. Shorter batteries have been developed for research in conditions ranging from attention-deficit hyperactivity disorder (ADHD) to Alzheimer's disease. There are four alternate forms of each test, and they are available in multiple languages. The average time required for completion of the MindStreams Global Assessment Battery is 45–60 minutes.

Test–retest reliability has been demonstrated over hours, weeks, and months, and correlation coefficients cited by NeuroTrax from multiple studies (e.g., Doniger & Simon, 2006; Schweiger et al., 2003) range from .52 for Memory to .83 for Motor Skills. The correlation cited for the Global Cognitive Score is .77. The low reliability on memory performance is a limitation in the battery.

Construct validity has been demonstrated in several studies, including a comparison of MindStreams and standard neuropsychological test performance in a cohort of 54 elderly subjects, some of whom were healthy and others met the criteria for Mild Cognitive Impairment (MCI; Dwolatzky et al., 2003). Correlations between the MindStreams Verbal Memory score and Wechsler Memory Scale-Third Edition (WMS III; Wechsler, 1997) Logical Memory I and II scores as well as Visual Reproduction II were in the range of .70, and the Nonverbal Memory score was correlated at the same level, with eight scores derived from the Rey Auditory Verbal Learning Test (RAVLT; Rey, 1961) and WMS III. Significant correlations were also reported between other MindStreams scores and scores on standard neuropsychological tests designed to measure the same constructs. Construct validity was also demonstrated in children and adolescents (Doniger & Simon, 2006; Schweiger et al., 2007) and in several clinical populations, including patients with movement disorders (Doniger et al., 2006b), multiple sclerosis (MS; Simon et al., 2006), schizophrenia (Ritsner et al., 2006), and Gaucher's disease (Elstein et al., 2005).

There are several versions of the MindStreams battery designed for research in ADHD, MCI and Alzheimer's disease, Parkinson's disease, and schizophrenia. There are also some limited data suggesting that MindStreams tests are sensitive to changes induced by alcohol (Jaffe et al., 2005) and to the effects of stimulant drugs in children with ADHD (Leitner et al., in press) and older patients with Parkinson's disease (Auriel et al., 2006), but these are small sample pilot studies.

In general, MindStreams appears to be a useful battery for the clinician assessing cognitive function and, particularly, attempting to distinguish among the healthy elderly those with MCI and those with mild dementia. In a study of 161 individuals with expert diagnoses (Doniger et al., 2005), the validity of the battery in making these distinctions was demonstrated convincingly. The value of the battery in some repeated-measures research will be limited by the availability of only four alternate forms. Of interest to clinicians will be the quality of the clinical report generated by MindStreams. In our view, the report is superior in detail and clarity to those produced by the other computerized batteries reviewed.

CogScreen

CogScreen (Kay, 1995) was developed twenty years ago in response to the Federal Aviation Administration's (FAA) need for an instrument that could detect subtle changes in cognitive function relative to poor pilot judgment or slow reaction time in critical flight situations. As such, CogScreen was intended to measure the underlying perceptual, cognitive, and information-processing abilities associated with flying.

In general, CogScreen is focused on measures of attention, concentration, information processing, immediate memory span, and working memory, as would be expected given the origin of the battery. CogScreen includes Backward Digit Span; Mathematical Reasoning; a Visual Sequence Comparison Task, in which the subject must identify two simultaneously presented alphanumeric strings as same or different; a Symbol Digit Coding Task (analogous to WAIS-R Digit Symbol), which also includes immediate and delayed incidental learning trials; a Matching-to-Sample Task involving presentation of a 4 × 4 grid of filled and empty cells, followed by a brief delay and presentation of two new matrices, one of which is identical to the original (requiring visuoperceptual speed, spatial processing, and visual working memory); the Manikin subtest, which presents a male human figure at different orientations holding a flag that the subject must identify as being in the left or right hand; a Divided Attention Task, combining visual monitoring and visual sequencing; an Auditory Sequence Comparison, requiring identification of two tonal sequences as the same or different (analogous to the Seashore Rhythm Test); a Pathfinder Test, similar to the Trail-Making Test; a Shifting Attention Test, requiring subjects to alter their responses dependent on changing rules (involving attribute identification, mental flexibility, sustained attention, deductive reasoning, response interference, and a variety of other cognitive skills); and Dual Task, which presents two tasks (visual-motor tracking and a visual memory span task for numbers) independently, then simultaneously. Most tests are available in 99 alternate forms.

The entire CogScreen battery requires about 45 minutes for administration, and all subject responses are input with a light pen or touch screen except for the tracking component of the Dual Task, which requires the subject to use the keyboard's arrow keys (Kane & Kay, 1992). The test battery runs on PCs with the Windows 98 or 2000 or XP operating system. There are now 10 different versions of CogScreen, which consist of different groupings of CogScreen subtests. These versions of CogScreen were developed for specific research applications. Test instructions are available in English, Spanish, French, and Russian. CogScreen is currently being used in North and South America, Europe, Asia, Africa, and Australia. Scoring is provided for response speed, accuracy, throughput (number of correct responses per minute), and, on certain tasks, process measures (e.g., impulsivity, perseverative errors). CogScreen also provides for entry of important demographic and performance-related variables (e.g., age, education, flight status, total flight hours logged, and nature of referral, such as alcohol-related or head injury). Norm-based reports are provided. The CogScreen U.S. aviator normative base includes 584 U.S. pilots screened for health status and alcohol and substance abuse. The CogScreen manual includes analysis of age effects, gender, education, and IQ on performance. Gender and education had minimal effects, whereas age and IQ reflected modest degrees of association with CogScreen performance (maximum age effect, 12.3% of the variance with any one measure; maximum IQ effect, 9% of variance).

Validity data are provided in terms of concurrent validity (correlation of CogScreen measures with related WAIS-R tasks, and correlations of CogScreen measures with specialized neuropsychological tasks such as the Wisconsin Card Sorting Test, PASAT, and Trail-Making Test). Several factor analytic studies of CogScreen have explored the factor structure of the battery (Taylor et al., 2000). Additional data are provided contrasting CogScreen performance of pilots selected from the normative base and matched on age and education with nonpilot controls and with patients with mild brain dysfunction. There were no significant differences between pilots and nonpilot controls; however, both groups performed at a superior level compared to the patient group. Additional data are provided on pilot groups with suspicion of alcohol abuse, pilots with questionable proficiency, pilots referred for evaluation of the impact of psychiatric impairment of cognitive function, and pilots with both suspected and confirmed neurologic disorders.

Although developed as a test to be used for the medical certification of pilots with psychiatric or neurological conditions, CogScreen has many other applications. In the aviation world, CogScreen is used for pilot selection by major airlines and by military organizations. The test has been proven to be a good predictor of pilot success. It is also used to periodically monitor the neurocognitive functioning of HIV seropositive pilots. In the area of biomedical research, CogScreen has been used to study the effect of environmental stressors on cognition (e.g., hypoxia and sleep deprivation), the neurocognitive effects of medical treatments (e.g., nasal CPAP treatment for sleep apnea), and for evaluating the adverse or beneficial cognitive effects of medications (e.g., the sedating effect of antihistamines).

The strength of CogScreen is in the area of attention and information-processing speed. The test includes sensitive measures of multitasking, working memory, and executive functions. On the other hand, it provides very limited testing of memory. For this reason, in clinical studies, CogScreen is often administered together with the Psychologix battery, described in later pages (e.g., Kay et al., 2006). CogScreen's use in "high stakes" testing (e.g., job selection, and as primary outcome measure in pharmaceutical and National Institutes of Health studies) is unparalleled. The validity of the test was affirmed by the National Transportation Safety Board (*Hoover v. FAA*). Furthermore, CogScreen results have been used to support drug claims made by pharmaceutical companies (e.g., the "nonsedating" label for Claritin).

Cognitive Drug Research (CDR) Battery

As the name implies, the CDR battery has been used almost exclusively in drug research. The origins of the battery date back to the late 1970s (Wesnes, 1977), and the core tests in the current battery were introduced in the mid-1980s. In the ensuing decades, the CDR was used in hundreds of clinical drug trials aimed at establishing efficacy or assessing unintended cognitive side effects of medication (e.g., Wesnes, 2003). Core tests in the CDR battery and the constructs they are intended to measure are given in Table 5–1.

Table 5–1. Core Tests in the CDR Battery and the Constructs They Measure

Attention
Simple Reaction Time
Choice Reaction Time
Working Memory
Articulatory Working Memory
Spatial Working Memory
Episodic secondary memory
Word Recall
Word Recognition
Picture Recognition

Each test is brief, usually one to three minutes, and subjects respond in testing by pressing "Yes" or "No" buttons. Fifty alternate forms are available for most tests, and very substantial data are available related to the reliability, validity, and utility of the system (Mohr et al., 1996). The reaction time tasks in the battery require no explanation, but the Articulatory Working Memory Test is based on the Sternberg procedure and involves the subject holding a short series, primarily through self-repetition, and then identifying whether digits presented subsequently are on the list being held. Both speed and accuracy are recorded on this and the other test of working memory. The Spatial Working Memory Test utilizes a three by three array of lights, said to be windows in a house. Four of the nine windows are randomly chosen and lighted, and on subsequent presentations individual windows are lighted and the subject must indicate whether or not that window was among the four initially lighted. The Episodic Secondary Memory tests utilize traditional recall (immediate and delayed) and recognition procedures.

Five factors have been identified among CDR outcome measures using Principal Components Analysis (Wesnes et al., 2000), and these, together with the measures that load on each factor, are shown in Table 5–2.

Although these are very simple and traditional tests, their use for more than two decades in many populations has generated a very significant body of data demonstrating, for example, sensitivity to and the existence of a distinctive test profile for aging, stroke, multiple sclerosis, chronic fatigue syndrome, diabetes, MCI, and many other conditions (Wesnes, 2003). CDR tests have been shown sensitive to the effects of a vast array of drugs and dietary supplements (e.g., Wesnes et al., 2000), but none of these findings has led to drug approval or a change in labeling, and, thus, questions related to specificity may arise.

The CDR battery is provided as a turnkey system consisting of a standard laptop computer with testing software, a proprietary USB key for data storage, and a response pad/box with

Table 5–2. CDR Outcome Measures Using Principal Components Analysis

Speed of Memory Processes
Picture Recognition Speed
Word Recognition Speed
Numeric Working Memory Speed
Spatial Working Memory Speed
Quality of Episodic Secondary Memory
Immediate Word Recall Accuracy
Delayed Word Recall Accuracy
Word Recognition Accuracy
Picture Recognition Accuracy
Power of Attention
Simple Reaction Time
Choice Reaction Time
Digit Vigilance Detection Speed
Continuity of Attention
Digit Vigilance Detection Accuracy
Choice Reaction Time Accuracy
Digit Vigilance False Alarms
Tracking Error
Quality of Working Memory
Numeric Working Memory Accuracy
Spatial Working Memory Accuracy

a separate USB connection. Tests can only be scored at CDR headquarters in England.

Cambridge Neuropsychology Test Automated Battery (CANTAB)

CANTAB, like CDR, comes from a British company that has developed very traditional tests and gathered quite a lot of data with them over the past two decades. Indeed, both companies claim to provide the world's most widely used computerized neuropsychological test battery. In the case of CANTAB, it is reported to be used in 50 countries, at 500 research institutes, with more than 300 publications. Also like CDR, CANTAB tests appear dated, although there is clear value in the databases that have been developed by both companies during the past 20 years. CANTAB, unlike CDR, has not focused primarily on drug development and appears less well suited to that task. CANTAB may be of greater interest to clinicians, however.

A current version of CANTAB, CANTAB eclipse, employs a touch screen as a primary response device, as well as a press pad for measuring reaction time. For many years before this version was introduced, CANTAB was used by researchers in aging, Alzheimer's disease, schizophrenia, depression, and many other conditions and disease states (e.g., Owen et al., 1990; Sahakian, 1990; Sahakian et al., 1988). It has also been shown sensitive to drug effects (e.g., Jones et al., 1992), although far more evidence of drug sensitivity has been shown for the CDR.

The CANTAB battery consists of 19 tests, beginning with two Motor Screening Tests, followed by four Visual Memory Tests. These are a Delayed Matching-to-Sample Test, including perceptual matching, immediate and delayed recall tasks in which the subject is shown a complex visual pattern, followed by four patterns from which he or she must choose the pattern first shown. This is followed by a Paired-Associates Learning Test, in which the subject must remember which patterns are associated with which positions on the touch screen. Still within the Visual Memory construct, a Pattern Recognition Memory Test is given, involving two-choice forced discriminations and a Spatial Recognition Memory Test, also involving two-choice forced discriminations. Moving on to a seemingly heterogeneous "Executive Function, Working Memory, and Planning" domain, the first test is the ID/ED Shift Test, which measures the ability to attend to a specific attribute of a complex visual stimulus and to shift the attribute when required. This is followed by the more colorfully named Stockings of Cambridge Test of spatial planning, based on the Tower of London Test. Next is the Spatial Span test, in which nine white squares appear in random positions on the screen and between two and nine of them are then lighted in different colors in random order. The subject must remember that order. Finally within this domain, the Spatial Working Memory Test involves a series of red boxes on the screen, some of which when touched, reveal a blue box. The object is to remember which red boxes one has touched and find the blue boxes without returning to a red box previously touched.

Performance in the Attention domain is assessed on CANTAB with several simple and complex reaction time tasks, a Matching-to-Sample Visual Search Test, and a Rapid Visual

Information-Processing test, in which the digits 2–9 are presented in a pseudorandom order at the rate of 100 per second. The subject must identify consecutive odd or even digits as quickly as possible.

There are two tests of the Semantic/Verbal Memory construct, the first of which is the Graded Naming Test, in which subjects must identify each of 30 black and white drawings of objects/animals presented in an increasing order of difficulty. The other test of the construct is the Verbal Recognition Memory test, in which subjects are presented with a list of 12 words, followed by free recall and forced choice recall. A final domain is Emotional Decision Making, and here there are two tests, the Affective Go–No Go Test, in which a subject must recognize the emotional valence of words that are presented on the screen, and the Cambridge Gambling Test, which is intended to assess decision-making and risk-taking behavior outside a learning context. This final test involves betting on the color of the box, among ten, that contains a token and appears appropriately named, but an odd measure of decision making in everyday life.

As noted, more than 300 studies have been reported using the CANTAB tests, and very substantial data on reliability and validity are available. Clinicians comparing MindStreams or CogState test presentation, scoring, and reporting with CANTAB may find the latter dated, particularly because CANTAB test computers must be purchased from the company and the tests cannot be downloaded and scored via the Internet. Also, the cost of the battery and a 10-year license is, as of this writing, US $14,000, and for these reasons the battery is used by few clinicians.

CogState

CogState tests (with one exception) are game-like and entirely lacking in face validity. There has been a serious effort in recent years to establish construct and criterion validity in the absence of face validity. Also, unlike CANTAB, CogState utilizes a technologically sophisticated platform for Internet testing. The company has slightly different batteries for clinical trials, academic studies, use in sports, testing by physicians (and presumably available to neuropsychologists), and use in the workplace. Tests in their Academic Battery, together with the cognitive domain said to be measured and the time required for administration, are listed in Table 5–3.

Table 5–3. Tests in CogState Academic Battery

CogState task	Cognitive domain	Time required (minutes)
Detection*	psychomotor function	2
Identification*	visual attention	2
One Card: learn*	visual learning	5
One Card: delayed recall	visual memory	
1.25		
One Word: Learn	verbal learning	5
One Word: Recall verbal memory		
1.25		
One Back*	working memory	2
Congruent reaction time*	visual attention	2
Monitoring*	Attention	2
Prediction*	executive function	5
Prediction: delayed recall*	visual memory	0.5
Associate learning*	Memory	5
International Shopping List Task (ISLT)	verbal learning	5
ISLT:delayed recall	verbal memory	1
Groton Maze Learning Task (GMLT)	executive function	10
GMLT: delayed recall	spatial memory	2
GMLT: reverse maze	executive function	2

A visual paired-associates learning and recall task, a set-shifting task, and a social–emotional cognition test have recently been added to the battery as well, although data on their reliability and validity have not been published to our knowledge.

CogState measures of test–retest reliabilities ranging from .67 for memory to .89 for psychomotor performance and also for attention have been reported, with playing cards utilized as the stimuli (Collie et al., 2003; Faletti et al., 2003).

Each of the tests above designated with an asterisk uses playing cards as stimuli, and, thus, the measures are remote from both traditional neuropsychological tests and clinical reality. Prior to the addition of the ISLT, the CogState battery was deficient in measures of verbal learning and memory, and this addition clearly strengthens the battery. Many of the tests in the battery also appear far too difficult for many clinical populations. On the other hand, the technology employed in testing and scoring is impressive, and the battery is well suited for use across languages and cultures. It is currently available in 15 languages and is easily translated into others. Also, the CogState battery has been used in a number of creative studies during the past 5 years, and more than 50 peer-reviewed articles have been published or are in press.

Maruff (In Press) has recently attempted to demonstrate construct validity and criterion validity by showing sensitivity to MCI, schizophrenia, and AIDS dementia. This study included only the tests in which playing cards are used as stimuli. One hundred and thirteen healthy young adults participated in the construct validity study, taking both CogState, a battery of standard neuropsychological tests, and a paired-associates memory test from the CANTAB computerized battery. Tests in the standard battery for psychomotor function were Trail Making A and Grooved Pegboard (dominant and nondominant). Attentional function was measured with the Digit Symbol Substitution test, cancellation task, and Trail Making Part B. The standard memory tests were the Paired-Associates Learning Test from the CANTAB, Benton Visual Retention, and delayed recall from the Rey Figure Test. Standard tests of executive function were the Spatial Working Memory Test strategy score, Spatial Span test, and the Tower of London Test from CANTAB; all tests are described in detail elsewhere (e.g., Lezak, 1995). In the psychomotor domain, correlations between standard and CogState measures ranged from .71 for Trails A to .32 for Grooved Pegboard (dominant). On attentional measures, correlations ranged from .67 on Digit Symbol and .56 on Trails B to .10 on the cancellation task. In the memory domain, all accuracy coefficients were highly significant ($p < .01$), ranging from .86 on the Rey Figure test and .85 on the CANTAB paired-associates test to .73 on Benton Visual Retention. Correlations were also highly significant on measures of executive function. On accuracy scores, the correlations were .86 with strategy on Spatial Working Memory, .69 with Spatial Span, and .67 with Tower of London performance. As to criterion validity, four separate studies reported by Maruff in this same paper demonstrate that CogState performance clearly distinguishes between comparable normals and individuals suffering from mild head injury resulting from auto accidents, MCI, schizophrenia, and AIDS dementia. The pattern of tests distinguishing subjects with each of the disorders from normals was as one would expect. For example, MCI deficits are limited largely to memory, while attentional and executive function deficits are also seen in schizophrenia. Other studies (e.g., Collie et al., 2002; Cysique et al., 2006) have also addressed the validity of Cogstate tests in distinguishing among these and other diagnostic groups. There are also data on the validity of the Groton Maze Learning Test (e.g., Pietrzak et al., 2007), but it is much more limited than that on tests using playing cards as stimuli.

CogState software can be downloaded over the Internet on to most PCs, and the primary response mechanism is the keyboard space bar. Scoring and reporting are done via the Internet.

The Automated Neuropsychological Assessment Metrics (ANAM)

ANAM is described as "a library of computerized tests and test batteries designed for a broad spectrum of clinical and research applications" (Reeves et al., 2007). The current version of ANAM, Version 4, is a Windows-based program that uses a mouse and keyboard as response input devices. There is also a "Web-enabled" and Palm-OS version of the test. The battery is available through C-Shop at the University of Oklahoma. The battery was originally developed by the Department of Defense and is similar to most of the early DOD Performance Assessment Batteries (PABs). There have been numerous iterations of the ANAM battery over the last 20 years. The test is highly configurable with respect to such parameters as inclusion or exclusion of subtests, number of items to include in a subtest, and interstimulus interval. The subtests included in ANAM are shown in Table 5–4.

In addition, there are Tapping, Tower of Hanoi, Stroop, and other standard cognitive tasks. Reeves described a "standard" or "default" version of ANAM that consists of 13 subtests. Although recognized, primarily, as a research tool, the creators of ANAM claim that it is now being developed as a clinical instrument.

A PubMed search of "ANAM" results in 25 citations between 1996 and 2007, with seven of these originating from the recent U.S. Army-sponsored supplement in the Archives of Clinical Neuropsychology. The articles in the supplement document the use of ANAM in "extreme environments," sports medicine, pharmacological studies, and in clinical populations. There are now very extensive norms available for active duty young military individuals ($N = 2371$). Prior publications document the use of ANAM in patients with systemic lupus erythematosus, hypothermia and Alzheimer's disease, multiple sclerosis, and traumatic brain injury, as well as individuals exposed to ionizing radiation. ANAM has also been used in evaluating the effects of a nicotine patch for treatment of Age-Associated Memory Impairment. In addition, studies have documented the reliability and validity of ANAM subtests.

Table 5–4. Subtests Included in ANAM

2-Choice Reaction Time
4-Choice Reaction Time
Code Substitution
Grammatical Reasoning
Logical Reasoning
Manikin
Matching Grids
Matching to Sample
Mathematical Processing
Memory Search
Running Memory (CPT)
Simple Reaction Time
Spatial Processing
Continuous Performance Test
Switching
Visual Vigilance

In spite of these developments, ANAM remains more of a "performance assessment battery" than a standardized test battery. The test is not widely used by either clinicians or investigators. Until recently, this DOD-funded test was in the "public domain." The test has been acquired by the University of Oklahoma, which now sells and licenses the battery.

ImPACT Test Battery

The ImPACT Test is a computer-administered test battery developed to assess concussion, primarily from sports-related injuries. The ImPACT 2005 is a Windows-based program that claims to measure response times with one-hundredth of a second resolution. The program is capable of creating an unlimited number of alternate forms. The test battery takes approximately 20 minutes and includes 6 "modules," which provide assessment of attention span, working memory, sustained and selective attention, response variability, nonverbal problem solving, and reaction time. There is a 20-minute delayed recall task for the word discrimination and design memory subtests. In addition to these two subtests, there are traditional measures of symbol digit substitution, a choice reaction task, a Sternberg three-letter recall task, and a visual working memory task. Results are immediately available upon completion of the exam in a well-designed report that compares the current scores to the subject's own baseline or to the normative database. The program can be installed on stand-alone PCs and does not require Internet connection for scoring or administration. The test generates a series of scores that are sensitive to head trauma, including five composite scores: Memory Composite (verbal), Memory Composite (visual), Visual Motor Speed, Reaction Time, and Impulse Control. The authors incorporated a Reliable Change Index for identifying meaningful changes in scores across administrations. The test is used by National Football League teams, Major League Baseball teams, and numerous colleges and universities. However, at present, the research base on the ImPACT test is almost entirely limited to sports injury studies (e.g., Miller et al., 2007).

Psychologix Computer-Simulated Everyday Memory Test Battery

The preceding batteries all share the feature of evaluating memory using traditional psychometric stimuli. By contrast, the Psychologix Computer-Simulated Everyday Memory Battery (previously known as the Memory Assessment Clinics Battery) is unique in that stimuli that are immediately relevant to everyday memory tasks are employed. The current battery represents the fifth generation of technology in a test development effort by Crook and colleagues, which began more that 25 years ago (e.g., Crook et al., 1979, 1980). From the outset, the goal was to simulate, in testing, tasks of everyday life that must be performed by virtually everyone, which are frequently affected by trauma, neurological disease, and developmental conditions. Tests were designed, in multiple alternative forms, using computer-imaging technology to simulate demands of everyday life. This heightened everyday realism was combined with traditional memory measurement paradigms such as selective reminding, signal detection, delayed nonmatching to sample, and associate learning.

Procedures include the *Name–Face Association Test* (Crook & West, 1990), in which live video recordings of persons introducing themselves are presented in different paradigms and both immediate and delayed recall are assessed. In the most frequently used paradigm, fourteen individuals appear on the screen, one after another, and each introduces himself/herself by saying, "Hi, I'm (First Name)." Each then reappears in a different order, and the task of the subject is to recall each name. On each recall trial the person to be remembered states the name of the city where he/she lives as a potential recall cue. Depending on the population, there are two or three such acquisition and immediate recall trials, followed by a delayed recall trial 30–40 minutes later. The city "cues" also provide the stimuli for an *Incidental Memory Test* (Crook et al., 1993), which assesses the subject's recall of the name of the city each person in *Name–Face Association* identifies as his/her home. Associative learning is also evaluated with a primarily nonvisual task, the *First–Last Names Test* (Youngjohn et al., 1991), which measures associate learning and recall of four to six paired first and last names over three to six trials (the subject must recall the first name associated with each last name). *Narrative Recall* measures the subject's ability to answer 25 factual, multiple-choice questions about a 6-minute television news broadcast (Crook et al., 1990a; Hill et al., 1989). *Selective Reminding* uses the paradigm devised by Buschke (1973) to evaluate learning and retention of 15 grocery items over five trials with a 30-minute delayed recall (Youngjohn et al., 1991). *The Misplaced Objects Test* (Crook et al., 1990b) is a visual–verbal associative task in which the subject "places" (by touching the touch-screen) 20 common objects (e.g., shoes, eyeglasses) in a 12-room house; 40 minutes later, the subject is given a first and a second chance at object location recall. Two measures are employed for facial recognition memory assessment (Crook & Larrabee, 1992). The first, *Recognition of Faces—Signal Detection*, employs signal detection procedures for evaluation of recognition memory, employing 156 facial photographs, with scores based on recognition over varying periods of time ranging from no delay to 1 minute, 3 minutes, and 5 minutes to a 40-minute delayed recognition period. The second, *Recognition of Faces—Delayed Nonmatching to Sample*, employs a delayed nonmatching to sample paradigm (Mishkin, 1978), in which the subject must identify, by touching the screen, the new facial photograph added to an array over 25 trials, each successive trial separated by an 8-second delay.

Working memory and attention are evaluated with the *Telephone Dialing Test* and *Reaction Time*. *Telephone Dialing* (West & Crook, 1990) requires the subject to dial 7- or 10-digit numbers on a representation of a telephone dialing pad on the computer screen after seeing them displayed on the screen. The test is also administered with an interference format, in which the subject, after dialing, hears either a ring or a busy signal. If the busy signal is heard, the subject must hang up, and then redial the telephone. *Reaction Time* (Crook et al., 1993) can be administered in one of two formats. First, reaction time can be measured under the single-task condition in which the subject must lift his or her finger off a computer-simulated (on the touchscreen) image of a gas pedal or brake pedal in response to a red or green traffic light. Both lift and travel (from gas to brake pedal or vice versa) reaction times are computed. Second, this task can be administered under the simultaneous processing task (divided attention) condition, where the subject must perform the gas pedal/brake pedal maneuvers while listening to a radio broadcast of road and weather conditions. In this administration, both lift and travel reaction times are computed—as well as the subject's recall of the radio broadcast information.

The Psychologix Battery has undergone extensive standardization and psychometric analysis. Crook and Larrabee (1988) and Tomer and colleagues (1994) factor analyzed a variety of scores from the battery and demonstrated factors of General Memory, Attentional Vigilance, Psychomotor Speed, and Basic Attention. The factor structure did not vary as a function of age, suggesting that although level of performance changed with age, the interrelationships of the tests did not. Hence, one can be assured that the tests are measuring the same constructs, regardless of the adult subject's age. In a second study, Larrabee and Crook (1989a) demonstrated a more varied factor structure, when First–Last Names and Selective Reminding were added to the battery. In this study, both verbal and visual factors emerged, in addition to an attentional factor and psychomotor speed factor. Concurrent validity was established by a combined factor analysis of the Psychologix Battery, WAIS Vocabulary, WMS-R Logical Memory and Paired-Associates Learning, and the Benton Visual Retention Test. Further evidence on validity is provided by many other studies. For example, Larrabee and Crook (1989b) reported cluster analyses that yielded a variety of everyday memory subtypes. Larrabee et al. (1991) demonstrated a significant canonical correlation of 0.528 between memory self-ratings (MAC-S; Crook & Larrabee, 1990) and factor scores from the Psychologix Battery. Youngjohn et al. (1992) reported a discriminant function analysis that correctly distinguished 88.39% of subjects with Alzheimer's disease from subjects with Age-Associated Memory Impairment. Also, Ivnik and colleagues (1993) demonstrated that the Reaction Time, Name–Face Association, and Incidental Memory procedures were highly sensitive to the cognitive effects of dominant temporal lobectomy. Psychologix tests have been shown highly reliable (Crook et al., 1992) and sensitive to the effects of drugs that both improve (e.g., Crook et al., 1991; Pfizer study) and impair (e.g., Kay et al., 2006; Nickelsen et al., 1999) cognition in many studies. Of greatest significance, because Psychologix tests are used almost exclusively in developing treatments for adult onset cognitive disorders, is the very high sensitivity of all tests to the effects of aging. For example, on the Name–Face Association Test (Crook & West, 1990; Crook et al., 1993), the decline in performance among healthy individuals is approximately 60% between ages 25 and 65. The effects of age on performance of Psychologix tests have been examined in more than 50 peer-reviewed publications, and individual differences that affect test performance have been examined in detail (e.g., West et al., 1992).

Crook et al. (1992) analyzed the equivalency of alternate forms of the various tests. At least six equivalent forms were found for Telephone Dialing, Name–Face Association, First–Last Name Memory, and Selective Reminding. Eight equivalent forms were found for Misplaced Objects and Recognition of Faces—Delayed Nonmatching to Sample.

American normative data for adults ages 18–90 range from 488 (TV News) to 2204 (Name-Face Memory), with sample sizes in the 1300 to 1900 range for most measures (Crook & West, 1990; Larrabee & Crook, 1994; West et al., 1992). Additional data are collected on over 500 persons with Alzheimer's disease and more than 2000 persons with Age-Associated Memory Impairment. The tests are available in several languages including English, a British (Anglicized) version, French, Italian, Swedish, Finnish, Danish, and German, and normative data are available in all these languages. A true random sample of the Italian population, comprising 1800 adults of all ages, is included in the normative database (Crook et al., 1993).

The principal current use of the Psychologix Battery is for evaluation of treatment effects in clinical trials of pharmacologic compounds, with potential benefit for ameliorating age-related memory disorders, and also in trials of drugs for the full range of medical indications where unwanted cognitive impairment may occur (e.g., Ferris et al., 2006; Kay et al., 2006). Although the current version of the Psychologix Battery is run on PCs with the Windows XP operating system and commercially available touchscreens, use of the battery requires specialized support, and, thus, it is not well suited to most clinical applications.

Summary

The test batteries reviewed in this chapter demonstrate the many advantages of computer-assisted cognitive evaluation, but it is important to exercise caution in the application of computers in clinical settings. Despite the technological sophistication of several available computerized memory tests, they do not all meet APA criteria for test instruments concerning reliability, validity, normative data, or with respect to their test manuals.

Nonetheless, significant progress has been made in the application of computers to memory evaluation in recent years. Continued growth can be expected in this area. There is a certain "technological seduction" concerning computerized assessment and remediation, but current as well as future applications of computers with cognitively impaired patients should be carefully considered in relation to the Division 40 guidelines for use of computers in evaluation and rehabilitation (Division 40 Task Force Report on Computer-Assisted Neuropsychological Evaluation, 1987; Matthews et al., 1991).

Future Directions

We believe that the use of computers to administer standard paper and pencil tests is a very early stage in the development of computerized neuropsychological testing. We believe that current technologies can be employed to provide highly realistic simulations of the cognitive tasks that must be performed in everyday life, on which developmental change or the effects of neurological disease or trauma are first noted. For example, we are now using sophisticated multiple-screen, computerized driving simulators that provide quite realistic graphics and representations of driving a car under a wide variety of circumstances. We have validated our driving simulator and shown it sensitive to drug effects (Kay et al., 2004), but beyond this technology lies the entire field of virtual reality. Future testing could allow subjects to enter an interactive world in which they would be called upon to perform a wide variety of cognitive tasks while their performance is precisely measured. Such development efforts are already under way (Rizzo, 2007). At present, we validate computerized tests against standard neuropsychological tests or in their ability to distinguish between groups based on age or disease states. Of course, our ultimate concern is with the individual's ability to function in a cognitively demanding environment, and, in our view, simulating that environment in cognitive testing is an important direction for the future.

References

Adams, K. M. (1986). Concepts and methods in the design of automata for neuropsychological test interpretation. In S. B. Filskov and T. J. Boll (Eds.), *Handbook of clinical neuropsychology* (vol. 2; pp. 561–576). New York: Wiley.

Adams, K. M., & Brown, G. C. (1986). The role of the computer in neuropsychological assessment. In I. Grant and K. M. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric disorders* (pp. 87–89). New York: Oxford.

Auriel, E., Hausdorff, J. M., Herman, T., Simon, E. S., & Giladi, N. (2006). Effects of methylphenidate on cognitive function and gait in patients with Parkinson's disease: A pilot study. *Clinical Neuropharmacology, 29*, 15–17.

Buschke, H. (1973). Selective reminding for analysis of memory and learning. *Journal of Verbal Learning and Verbal Behavior, 12*, 543–549.

Collie, A., Maruff, P., & Currie, J. (2002). Behavioral characterization of mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology, 24*, 720–733.

Collie, A., Maruff, P., Darby, D. G., & McStephen, M. (2003). The effects of practice on the cognitive test performance of neurologically normal individuals assessed at brief test–retest intervals. *Journal* *of the International Neuropsychological Society, 9,* 419–428.

Crook, T. H., Ferris, S., & McCarthy, M. (1979). The Misplaced-Objects Task: A brief test for memory dysfunction in the aged. *Journal of the American Geriatrics Society, 27*, 284–287.

Crook, T. H., Ferris, S. H., McCarthy, M., & Rae, D. (1980). Utility of digit recall tasks for assessing memory in the aged. *Journal of Consulting and Clinical Psychology, 48*, 228–233.

Crook, T.H., & Larrabee, G.J. (1988). Interrelationships among everyday memory tests: Stability of factor structure with age. *Neuropsychology, 2*, 1–12.

Crook, T. H., & Larrabee, G. J. (1990). A self-rating scale for evaluating memory in everyday life. *Psychology and Aging, 5*, 48–57.

Crook, T. H., & Larrabee, G. J. (1992). Changes in facial recognition memory across the adult life span. *Journal of Gerontology, 47*, 138–141.

Crook, T. H., Larrabee, G. J., & Youngjohn, J. R. (1993). Age and incidental recall for a simulated everyday memory task. *Journal of Gerontology, 48*, 45–47.

Crook, T. H., Tinklenberg, J., Yesavage, J., Petrie, W., Nunzi, M. G., & Massari, D. C. (1991). Effects of phosphatidylserine in Age Associated Memory Impairment. *Neurology, 41*, 644–649.

Crook, T. H., & West, R. L. (1990). Name recall performance across the adult life span. *British Journal of Psychology, 81*, 335–349.

Crook, T. H., West, R. L., & Larrabee, G. J. (1993). The Driving-Reaction Time Test: Assessing age declines in dual-task performance. *Developmental Neuropsychology, 9*, 31–39.

Crook, T. H., Youngjohn, J. R., & Larrabee, C. J. (1990a). The TV News Test: A new measure of everyday memory for prose. *Neuropsychology, 4*, 135–145.

Crook, T. H., Youngjohn, J. R., & Larrabee, C. J. (1990b). The Misplaced Objects Test: A measure of everyday visual memory. *Journal of Clinical and Experimental Neuropsychology, 12*, 819–833.

Crook, T. H., Youngjohn, J. R., & Larrabee, C. J. (1992). Multiple equivalent test forms in a computerized everyday memory battery. *Archives of Clinical Neuropsychology, 7*, 221–232.

Cysique, L. A., Maruff, P., Darby, D., & Brew, B. J. (2006). The assessment of cognitive function in advanced HIV-1 infection and AIDS dementia complex using a new computerized cognitive test battery. *Archives of Clinical Neuropsychology, 21*, 185–194.

Division 40 Task Force Report on Computer Assisted Neuropsychological Evaluation. (1987). *Clinical Neuropsychologist, 2*, 161–184.

Doniger, G. M., Dwolatzky, T., Zucker, D. M., Chertkow, H., Crystal, H., Schweiger, A., et al. (2006a). Computerized cognitive testing battery identifies MCI and mild dementia even in the presence of depressive symptoms. *American Journal of Alzheimer's Disease and Other Dementias, 21,* 28–36. AQ3

Doniger, G. M., Okun, M. S., Simon, E. S., Rodriguez, R. L., Jacobson, C. E., Weiss, D., et al. (2006b). Validation of a computerized neuropsychological assessment (Mindstreams) in movement disorders: Interim analysis. *20th Annual Symposia on the Etiology, Pathogenesis, and Treatment of Parkinson's Disease and Other Movement Disorders.* AQ4

Doniger, G. M., & Simon, E. S. (2006). *Construct Validity of Mindstreams® Comparison with Paper-Based Tests.* Internal Document. New York: NeuroTrax Corporation.

Doniger, G. M., Zucker, D. M., Schweiger, A., Dwolatzky, T., & Simon, E. S. (2005). Towards practical cognitive assessment for detection of early dementia: A 30-minute computerized battery discriminates as well as longer testing. *Current Alzheimer Research, 2,* 117–124.

Dwolatzky, T., Whitehead, V., Doniger, G. M., Simon, E. S., Schweiger, A., Jaffe, D, et al. (2003). Validity of a novel computerized cognitive battery for mild cognitive impairment. *BMC Geriatrics, 3,* 4.

Dwolatzky, T., Whitehead, V., Doniger, G. M., Simon, E. S., Schweiger, A., Jaffe, D., et al. (2004). Validity of the Mindstreams computerized cognitive battery. *Journal of Molecular Neuroscience, 24,* 33–44.

Elstein, D., Guedalia, J., Doniger, G. M., Simon, E. S., Antebi, V., Arnon, Y., et al.. (2005). Computerized cognitive testing in patients with type I Gaucher disease: Effects of enzyme replacement and substrate reduction. *Genetics in Medicine, 7,* 124–130.

Elwood, R. W. (2001). MicroCog: Assessment of cognitive functioning. *Neuropsychology Review, 11*, 89–100.

Falleti, M. G., Maruff, P., Collie, A., Darby, D. G., & McStephen, M. (2003). Qualitative similarities in cognitive impairment associated with 24 h of sustained wakefulness and a blood alcohol concentration of 0.05%. *Journal of Sleep Research, 12*, 265–274.

Ferris, S., Schneider, L., Farmer, M., Kay, G., & Crook, T. (2006). A double-blind, placebo-controlled trial of memantine in age-associated memory impairment (memantine in AAMI). *International Journal of Geriatric Psychiatry, 21*, 1–8.

Helmes, E., & Miller, M. (2006). A comparison of MicroCog and the Wechsler Memory Scale (3rd ed.) in older adults. *Applied Neuropsychology, 13*, 28–33.

Hill, R. D., Crook, T. H., Zadeik, A., Sheikh, J., & Yesavage, J. (1989). The effects of age on recall of information from a simulated television news broadcast. *Educational Gerontology, 5*, 607–613.

Ivnik, R. J., Malec, J. F., Sharbrough, F. W., Cascino, G. D., Hirschorn, K. A., Crook, T. H., et al. (1993). Traditional and computerized assessment procedures applied to the evaluation of memory change after temporal lobectomy. *Archives of Clinical Neuropsychology, 8,* 69–81.

Jaffe, D., Doniger, G. M., Simon, E. S., & Neumark, Y. (2005). Computerized cognitive test performance varies with acute alcohol consumption. *158th Annual Meeting of the American Psychiatric Association*. August 18–21, Washington, DC.

Jones, G., Sahakian, B. J., Levy, R., Warburton, D. M., & Gray, J. (1992). Effects of acute subcutaneous nicotine on attention, information processing and short-term memory in Alzheimer's disease. *Psychopharmacology, 108*, 485–494.

Kane, R. L. (1995). MicroCog: A review. *Bulletin of the National Academy of Neuropsychology, 11*, 13–16.

Kane, R. L., & Kay, G. G. (1992). Computerized assessment in neuropsychology: A review of tests and test batteries. *Neuropsychology Review, 3,* 1–117.

Kapur, N. (1988). *Memory disorders in clinical practice.* London: Butterworths.

Kay, G. G. (1995). *CogScreen-Aeromedical Edition. Professional Manual*. Odessa, FL: Psychological Assessment Resources.

Kay, G. G., Crook, T., Rekeda, L., Lima, R., Ebinger, U., Arguinzoniz, M., et al. (2006). Differential effects of the antimuscarinic agents darifenacin and oxybutynin ER on memory in older subjects. *European Association of Urology, 50,* 317–326.

Kay, G. G., Pakull, B., Clark, T. M., Sea, D., Mays, D. A., & Tulloch, S. J. (2004). The effect of Adderall XR® treatment on driving performance in young adults with ADHD. *Paper presented at 17th Annual US Psychiatric and Mental Health Congress*, San Diego, CA, November 2004.

Larrabee, G. J., & Crook, T. H. (1989a). Dimensions of everyday memory in Age Associated Memory Impairment. *Psychological Assessment, 1,* 92–97.

Larrabee, G. J., & Crook, T. H. (1989b). Performance subtypes of everyday memory function. *Developmental Neuropsychology, 5,* 267–283.

Larrabee, G. J., & Crook, T. H. (1991). Computerized memory testing in clinical trials. In E. Mohr and P. Brouwers (Eds.), *Handbook of clinical trials: The neurobehavioral approach* (pp. 293–306). Amsterdam: Swets & Zeitlinger.

Larrabee, G. J., & Crook, T. H. (1994). Estimated prevalence of Age-Associated Memory Impairment derived from standardized tests of memory function. *International Psychogeriatrics, 6*, 95–104.

Larrabee, G. J., & Crook, T. H. (1996). Computers and memory. In I. Grant and K. M. Adams (Eds.). *Neuropsychological assessment of neuropsychiatric disorders (Second Edition)* (pp. 102–117). New York: Oxford University Press.

Ledbetter, M., & Hurley, S. (1994). A construct validation study of the attention and memory domains on the Assessment of Cognitive Functioning Test. *Archives of Clinical Neuropsychology, 9,* 154.

Leitner, Y., Doniger, G. M., Barak, R., Simon, E. S., & Hausdorff, J. M. (In Press). A novel multi-domain computerized cognitive assessment for attention deficit hyperactivity disorder: Evidence for widespread and circumscribed cognitive deficits. *Journal of Child Neurology*.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd. ed.). New York: Oxford University Press.

Matthews, C. G., Harley, J. P., & Malec, J. F. (1991). Guidelines for computer-assisted neuropsychological rehabilitation and cognitive remediation. *Clinical Neuropsychologist, 5,* 319.

Maruff, P. (In Press). Validity of the CogState computerized cognitive test battery: Sensitivity to mild traumatic brain injury, mild cognitive impairment, schizophrenia and AIDS dementia complex. *The Clinical Neuropsychologist.*

Miller, J. R., Adamson, G. J., Pink, M. M., & Sweet, J. C. (2007). Comparison of preseason, midseason, and postseason neurocognitive scores in uninjured collegiate football players. *American Journal of Sports Medicine, 35*(8), 1284–1288.

Mishkin, M. (1978). Memory in monkeys severely impaired by combined but not by separate removal of amygdale and hippocampus. *Nature, 273*, 297–298.

Mittenberg, W., DiGiulio, D. V., Perrin, S., & Bass, A. E. (1992). Symptoms following mild head injury: expectation as etiology. *Journal of Neurology, Neurosurgery, and Psychiatry, 55*, 200–204.

Mohr, E., Walker, D., Randolph, C., Sampson, M., & Mendis, T. (1996). The utility of clinical trial batteries in the measurement of Alzheimer's and Huntington's dementia. *International Psychogeriatrics, 8,* 397–411.

Morris, R. G., Evenden, J. L., Sahakian, B. J., & Robbins, T. W. (1987). Computer-aided assessment of dementia: comparative studies of neuropsychological deficits in Alzheimer-type dementia and Parkinson's disease. In S. M. Stahl, S. D. Iversen, and T. W. Robbins (Eds.). *Cognitive neurochemistry* (pp. 21–36). New York: Oxford University Press.

Nickelsen, T., Lufkin, E. G., Riggs, B. L., Cox, D. A., & Crook, T. H. (1999). Raloxifene hydrochloride, a selective estrogen receptor modulator: Safety assessment of effects on cognitive function and mood in postmenopausal women. *Neuroimmunoendocrinology, 24,* 115–128.

Owen, A. M., Downes, J. J., Sahakian, B. J., Polkey, C. E., & Robbins, T. W. (1990). Planning and spatial working memory following frontal lobe lesions in man. *Neuropsychologia, 28,* 1021–1034.

Pietrzak, R. H., Cohen, H., & Snyder, P. J. (2007). Spatial learning efficiency and error monitoring in normal aging: An investigation using a novel hidden maze learning test. *Archives of Clinical Neuropsychology, 22*, 235–245.

Powell, D. H., Kaplan, E. F., Whitla, D., Weintraub, S., Catlin, R., & Funkenstein, H. H. (1993). *MicroCog. Assessment of Cognitive Functioning. Manual.* Orlando, FL: The Psychological Corporation.

Raymond, P., Hinton-Bayre, A., Radel, M., Ray, M., & Marsh, N. (2006). Test–rest norms and reliable change indices for the MicroCog battery in a healthy community population over 50 years of age. *The Clinical Neuropsychologist, 20,* 261–270.

Reeves, D. L., Winter, K. P., Bleiberg, J., & Kane, R. L. (2007). ANAM® Genogram: Historical perspectives, description, and current endeavors. *Archives of Clinical Neuropsychology, 22S*, S15–S37.

Rey, A. (1961). *L'examen clinique en psychologie.* Paris: Presses Universitaires de France.

Ritsner, M. S., Blumenkrantz, H., Dubinsky, T., & Dwolatzky, T. (2006). The detection of neurocognitive decline in schizophrenia using the Mindstreams Computerized Cognitive Test Battery. *Schizophrenia Research, 82,* 39–49.

Rizzo, A. (2007). T13: Virtual reality in mental health and rehabilitation: An overview. *HCI International 22–27 July 2007, Beijing, China*.

Sahakian B. J. (1990). Computerized assessment of neuropsychological function in Alzheimer's disease and Parkinson's disease. *International Journal of Geriatric Psychiatry, 5,* 211–213.

Sahakian, B. J., Morris, R. G., Evenden, J. L., Heald, A., Levy, R., Philpot, M. P., et al. (1988). A comparative study of visuospatial memory and learning in Alzheimer-type dementia and Parkinson's disease. *Brain, 111*, 695–718.

Schweiger, A., Abramovitch, A., Doniger, G. M., & Simon, E. S. (2007). A clinical construct validity study of a novel computerized battery for the diagnosis of ADHD in young adults. *Journal of Clinical and Experimental Neuropsychology, 29,* 100–111.

Schweiger, A., Doniger, G. M., Dwolatzky, T., Jaffe, D., & Simon, E. S. (2003). Reliability of a novel computerized neuropsychological battery for mild cognitive impairment. *Acta Neuropsychologica, 1,* 407–413

Seltzer, B., Zolnouni, P., Nunez, M., Goldman, R., Kumar, D., Ieni, J., et al. (2004). Efficacy of donepizil in early stage Alzheimer's disease. *Neurology, 61,* 1852–1856.

AQ5

Simon, E. S., Harel, Y., Appleboim, N., Doniger, G. M., Lavie, M., & Achiron, A. (2006). Validation of the Mindstreams Computerized Cognitive Battery in multiple sclerosis. *Neurology, 66,* A239.

Suhr, J. A., & Gunstad, J. (2002). "Diagnosis threat": The effect of negative expectations on cognitive performance in head injury. *Journal of Clinical and Experimental Neuropsychology, 24*, 448–457.

Taylor, J. L., O'Hara, R., Mumenthaler, M. S., & Yesavage, J. A. (2000). Relationship of CogScreen-AE to Flight Simulator Performance and Pilot Age. *Aviation, Space, and Environmental Medicine, 71*, 373–380.

Tomer, A., Larrabee, G. J., & Crook, T. H. (1994). Structure of everyday memory in adults with Age-Associated Memory Impairment. *Psychology and Aging, 9*, 606–615.

Wechsler, D. A. (1987). *Wechsler Memory Scale Revised: Manual.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. A. (1997). *Wechsler Memory Scale; 3rd Edition (WMS-III).* San Antonio, TX: The Psychological Corporation.

Wesnes, K. A. (1977). The effects of psychotropic drugs on human behaviour. *Modern Problems of Pharmacopsychiatry, 12,* 37–58.

Wesnes, K. A. (2003). The cognitive drug research computerised assessment system: Application to clinical trials. In P. De Deyn, E. Thiery, and R. D'Hooge (Eds.), *Memory: Basic concepts, disorders and treatment* (pp. 453–472).Leuven, Belgium: ACCO.

Wesnes, K. A., Simpson, P. M., & Christmas, L. (1987). The assessment of human information processing abilities in psychopharmacology. In I. Hindmarch and P. D. Stonier (Eds.), *Human psychopharmacology: Measures and Methods, 1*, 79–92.

Wesnes, K. A., Ward, T., McGinty, A., & Petrini, O. (2000). The memory enhancing effects of a Ginko biloba/Panax ginseng combination in healthy middle-aged volunteers. *Psychopharmacology, 152*, 353–361.

West, R. L., & Crook, T. H. (1990). Age differences in everyday memory: Laboratory analogues of telephone number recall. *Psychology and Aging, 5,* 520–529.

West, R. L., Crook, T. H., & Barron, K. L. (1992). Everyday memory performance across the life span: The effects of age and noncognitive individual difference factors. *Psychology and Aging, 7,* 72–82.

Youngjohn, J. R., Larrabee, G. J., & Crook, T. H. (1991). First–last names and the grocery list selective reminding tests: Two computerized measures of everyday verbal learning. *Archives of Clinical Neuropsychology, 6,* 287–300.

Youngjohn, J. R., Larrabee, G. J., & Crook, T. H. (1992). Test–retest reliability of computerized, everyday memory measures and traditional memory tests. *Clinical Neuropsychologist, 6,* 276–286.

AQ1: Please provide complete reference details for cross-reference 'Maruff, 2007' in the reference list.
AQ2: Please provide complete reference details for cross-reference 'Larrabee et al., 1991' in the reference list.
AQ3: Please confirm whether the insertion of 'a' to the year of publication is correct, as there are two entries with the same author name and same year of publication.
AQ4: Please provide the location where the symposia took place.
AQ5: Please provide cross-reference for reference 'Seltzer et al., 2004' in the text.