Intern. J. Neuroscience, 115:1549–1567, 2005

ISSN: 0020-7454 / 1543-5245 online
DOI: 10.1080/00207450590957890

Taylor & Francis
Taylor & Francis Group

PRELIMINARY VALIDITY OF "INTEGNEURO™": A NEW COMPUTERIZED BATTERY OF NEUROCOGNITIVE TESTS

ROBERT H. PAUL
JEFFREY LAWRENCE

Brown Medical School
Department of Psychiatry
Centers for Behavioral and Preventive Medicine
Providence, Rhode Island, USA

LEANNE M. WILLIAMS

Department of Psychological Medicine
University of Sydney
Sydney, Australia

and

The Brain Dynamics Centre Westmead Hospital
New South Wales, Australia

and

School of Psychology
University of Sydney
Sydney, Australia

C. CLARK RICHARD

Flinders University
Adelaide, Australia

Received 27 November 2004.
Address correspondence to Robert H. Paul, Brown Medical School, Coro Building, Behavioral Medical Clinic, 1 Hoppin St., Providence, RI 02903, USA. E-mail: rpaul@lifespan.org

NICHOLAS COOPER

The Brain Resource International Database
Brain Resource Company
New South Wales, Australia

EVIAN GORDON

The Brain Resource International Database
Brain Resource Company
New South Wales, Australia

and

Department of Psychological Medicine
University of Sydney
Sydney, Australia.

and

The Brain Dynamics Centre Westmead Hospital
New South Wales, Australia

The purpose of this study was to examine the preliminary validity of a newly developed battery of computerized cognitive measures, IntegNeuroTM. This standardized and semi-automated computerized battery assesses sensori-motor function, attention, new learning and memory, language fluency, executive function, and estimated intelligence. A total of 50 healthy individuals (aged 22–80 years) were included in the study. Correlational analyses revealed highly significant associations between the two cognitive batteries. These results support the use of IntegNeuroTM as a computerized cognitive system. Additional studies are needed to examine the clinical utility of the battery.

Keywords assessment, computerized testing, neuropsychology, validity

INTRODUCTION

Computerized neuropsychological assessment has increasingly received recognition as a valuable research and clinical tool. The American Psychological Association (APA) recognized the value of computerized psychological testing and published guidelines in 1987 (APA, 1987) to assist in the development and interpretation of computerized test results. In this publication, the APA identified six major benefits of computerized assessment including: (1)

automated data collection and storage, (2) greater efficiency of use, (3) release of the clinician from test administration to focus on treatment, (4) greater sense of mastery and control for the client, (5) reduced negative self-evaluation among clients that experience difficulty on the computer, and (6) greater ability to measure aspects of performance not possible through traditional means such as latency, strength, and variability in response patterns.

Despite their promise, initial efforts to computerize cognitive tests focused mainly on individual measures, and the validity of these outcomes varied (for review see Kane & Kay, 1992). Further, many early versions of computerized assessment suffered from poor visual graphics, inadequate sound quality, and inconsistencies in recording of responses. These limitations have been overcome with developments in computing hardware and software. Computer methods now exist that allow for accurate and reliable timing of stimulus presentation, response recording, and multidimensional display of information.

A number of computerized cognitive batteries have been developed in recent years on the heels of the advancements in computerized technology. Three computerized batteries frequently cited in the literature include the Cambridge Automated Neuropsychological Test Battery (CANTAB; Morris et al., 1986), the MicroCog (Devivo et al., 1999), and the Neurobehavioral Evaluation System (Baker et al., 1985). These computerized batteries have provided significant contributions to the research literature. There are, however, aspects of these batteries that restrict their utility, including the absence of language measures on the CANTAB and NES 3, and the assessment of cognitive constructs that differ from standardized clinical assessment (e.g., MicroCog; Elwood, 2001). As such, there is a need for the development of cognitive programs that capitalize on the advancements of computing technology to allow assessment of standard cognitive skills including language and verbal memory.

IntegNeuro™ (Brain Resource Company, Ltd) is a newly developed computerized battery that consists of an automated stimulus presentation protocol, and response recording involving touch-screen and voice recording software. The tests were based on existing paradigms known to be sensitive to brain dysfunction. Attractive features of the battery include standardized instructions using both auditory explanations and visual examples, practice trials prior to test trials, and semi-automated scoring procedures. In addition, the battery include both language and nonlanguage paradigms. The battery was developed by a consortium of scientists involved in the first standardized international brain database (Gordon, 2005). Test-retest reliability for each of the cognitive tests is acceptable for all measures (Williams et al., current volume).

The purpose of the present study was to examine the preliminary validity of IntegNeuro™, via examination of the relationships between performances on IntegNeuro™ and performances on previously developed common tests that tap the same cognitive skills. It was predicted that the individual measures of IntegNeuro™ would correlate strongly with analogous standard measures of cognitive function.

METHODS

Participants

A total of 50 healthy adults (25 females and 25 males, age 22–80) completed the IntegNeuro™ and the standard neuropsychological batteries. Exclusion criteria included any mental or physical condition with the potential to influence cognitive performance, including a personal history of mental illness, physical brain injury, neurological disorder, genetic disorder, or other medical condition (hypertension, diabetes, cardiac disease, thyroid disease), and/or a personal history of drug or alcohol addiction. The computerized Composite International Diagnostic Interview (CIDI, WHO, 1993) was also used to exclude participants if they had a family history of Attention Deficit Hyperactivity Disorder, Schizophrenia, or Bipolar Disorder. The CIDI consists of computer-aided self report endorsement of psychiatric symptoms and diagnoses are determined according to criteria defined in the Diagnostic and Statistical Manual—IV (APA, 1994). The CIDI is both valid and reliability for multi-site and international use (Wittchen, 1991; Wittchen et al., 1999).

Web-based questionnaires were used to acquire demographic data including age, sex, years of education and current mood state in terms of depression, anxiety, and stress (assessed using an abbreviated version of the Depression Anxiety Stress Scale; DASS; Lovibond & Lovibond, 1995). These demographic data are presented in Table 1 for the full sample. All participants voluntarily signed a written informed consent form to participate in the database, according to local Institutional Review Boards.

Procedure

IntegNeuro™ was administered in a sound-attenuated testing room, with participants seated in front of a touch-screen computer (NEC MultiSync LCD 1530V). The cognitive tests were administered using standardized task instructions (via headphones and visual screen display), and the touch-screen

Table 1. Demographic data for the sample

	Mean (SD)
Age	49.2 (18.3)
Education	15.3 (2.0)
DASS depression*	1.9 (1.9)
DASS anxiety*	1.0 (1.3)

*Note that these scores are below recommended cutoffs for clinical significance (i.e., < 4.0; Lovibond & Lovibond, 1995).

computer was used to record nonverbal responses. Examples of the system are provided in Figure 1. The paper-and-pencil tests were administered by a highly trained psychometrician according to standardized manuals. In one half the cases (n = 25), the IntegNeuro™ battery was administered at the first visit, and four weeks later the previously developed standard measures were administered at a second visit. The other half of the cases (25 individuals) first completed the standard measures and subsequently completed IntegNeuro™.

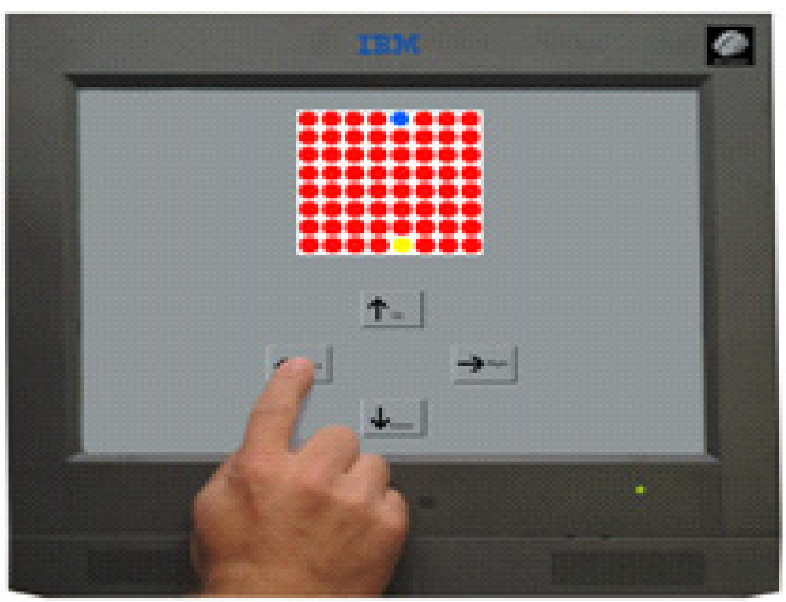

Figure 1. Visual example of the Maze task on IntegNeuro™.

The order of administration (IntegNeuro™ vs. standard measures) was determined by random assignment to avoid an order effect. Each measure of IntegNeuro™ included a practice trial prior to the test trial. Individuals were required to pass the practice trial accurately before completing the test trial. In the event that an individual was unable to perform the practice trial without error, the individual test was terminated and the individual was automatically forwarded to the next test in the battery. In the present study, all participants were capable of passing the practice trials.

Cognitive Tests

IntegNeuro. The battery of IntegNeuro™ tests tapped the following domains of cognitive function: sensori-motor, verbal and language, memory, executive planning, and attention. Scoring of responses was conducted using an automated software program for most tests, and by hand-scoring for .wav files. Hand scoring was required for the two language tests and the verbal memory test. Trained research assistants conducted the hand scoring of the .wav files and oversight was implemented to monitor accuracy. The measures in each of these domains are described in the following.

Sensori-Motor Domains. (i) Simple motor tapping task: Participants were required to tap a circle on the touch-screen with their index finger, as fast as possible for 60 s. The dependent variable was total number of taps with the dominant hand. (ii) Choice reaction time task: Participants were required to attend to the computer screen as one of four target circles was illuminated in pseudo random sequence over a series of trials. For each trial, the subject was required to place their index finger in preparation on a start circle displayed on the touchscreen. On each trial, the subject then had to touch the illuminated circle as quickly as possible following presentation. Twenty trials were administered with a random delay between trials of 2–4 s. The dependent variable was the mean reaction time across trials.

Attention Domain. (i) Span of Visual Memory: This test is a computerized adaptation of the Spatial Span test from the Wechsler Memory Scale (III; Wechsler, 1999b). Participants were presented with squares arranged in a random pattern on the computer screen. The squares were highlighted in a sequential order on each trial. Participants were required to repeat the order in which the squares were highlighted by touching the squares with their forefinger. Both forward and reverse trials are conducted. The total correct was the dependent variable. (ii) Digit Span: Participants were presented with a series of digits on the touchscreen (e.g., 4, 2, 7, etc., 500 ms presentation), separated by a 1-s interval. The subject was then immediately asked to enter the digits on a numeric keypad on the touch-screen. In the first part of the test, subjects were required to recall the digits in forward order (Digits forwards); in the second part, they were required to recall them in reverse order (Digits backwards). In each part, the number of digits in each sequence was gradually increased from 3 to 9, with two sequences at each level. The dependent measure for each part was the maximum number of digits the subject recalled without error. (iii) Continuous Performance task: To tap sustained attention, a series of letters (B, C, D, or G) were presented to the subject on the computer screen (for 200 ms), separated by an interval of 2.5 s. If the same letter appeared twice in a row, the subject was asked to press buttons with the index finger of each hand. Speed and accuracy of response were equally stressed in the task instructions. There were 125 stimuli presented in total, 85 being non-target letters, and 20 being target letters (i.e., repetitions of the previous letter). The dependent variables were the number of errors of omissions and false positives. (iv) Switching of attention task: this test is a computerized adaptation of the Trail Making test (Reitan, 1958). It consists of two parts. In the first part, the subject was presented with a pattern of 25 numbers in circles and asked to touch them in ascending numerical sequence (i.e., 1, 2, 3, . . .). As each number is touched in correct order, a line is drawn automatically to connect it to the preceding number in the sequence. This allowed the subject to visualize the path touched. This task tests psychomotor speed and the basic ability to hold attention on a simple task. The second part of the test is described in what follows. The dependent variable was time to completion.

Executive Function Domain. (i) Switching of attention task; part 2: In the second part of this task, the subject was presented with a pattern of 13 numbers (1–13) and 12 letters (A–L) on the screen and was required to touch numbers and letters alternatively in ascending sequence (i.e., 1, A, 2, B, 3, C, . . .). This part is harder than the first part and reflects the requirement to switch attention between mental tasks, in this case number and letter sequence checking, and thereby alternate between the respective mental sets induced. The dependent variable was time to completion. (ii) Verbal Interference: This task taps the ability to inhibit automatic and irrelevant responses and has similarities to the Stroop task (Golden, 1978). The subject was presented with colored words presented serially, one at a time. Each word was drawn from the following set of lower case words: red, yellow, green, and blue. The color of each word is drawn from the following set of colors: red, yellow, green, and blue. Below each colored word is a response pad with the four possible words displayed in black and in fixed format. The test has two parts. In part 1, the subject is required to identify the name of each word as quickly as possible after it is presented on the screen. In part 2, the subject is required to name the color of each word as quickly as possible. Each part lasts for 1 min. Responses are made on the screen by touching the appropriate word on the response pad. The dependent variable in each part was the number of words correctly identified. (iii) Maze Task: This task was a computerized adaptation of the Austin Maze (Walsh, 1985). The subject was presented with a grid (8 × 8 matrix) of circles on the computer screen. The object of the task was to identify the hidden path through the grid, from the beginning point at the bottom of the grid to the end point at the top. The subject was able to navigate around the grid by pressing arrow keys (up, down, left, right). A total of 24 consecutive correct moves were required to complete the maze. The subject was presented with one tone (and a red cross at the bottom of the screen) if they made an incorrect move, and a different tone (and a green tick at the bottom of the screen) if they made a correct move. The purpose of the task was therefore to assess how quickly the subject learned the route through the maze and their ability to remember that route. Only one maze was presented across trials, and the test ended when the subject completed the maze twice without error or after 10 min had elapsed. The dependent variable was the total maze time. It should be noted that although this measure is identified as a measure of executive function, the requirement to retain the maze in memory for two successive trials introduces an added memory component to the task, and therefore this measure taps more cognitive domains than pure executive function.

Language Domain. (i) Letter Fluency: Participants were required to generate by speech words that began with the letters F, A, and S. 60 s were allowed for each letter and proper nouns were not allowed. Responses were recorded via the microphone and hand scored. Intrusive or perseverative responses were not included in the total number correct. The total number of correct words generated across the three trials was the dependent measure. (ii) Animal Fluency: Participants were required to name animals as quickly as possible for 60 s. Intrusions and perseverative responses were not allowed. Total correct served as the dependent variable.

Memory Domain. (i) Verbal List-learning: The participants were read a list of 12 words, which they were asked to memorize. The list contained 12 concrete words from the English language. Words were closely matched on concreteness, number of letters, and frequency. The list was presented orally 4 times (and received by the subject using headphones). On each of the 4 trials, the subject was required to recall as many words as possible by speaking directly into the attached microphone. The subject was then presented with a list of distracter words and asked to recall them after presentation. Immediately following this, the subject was then asked to recall the 12 words from the original list (short-delay recall trial). A long delayed recall trial was completed approximately 20 min later after a number of intervening tasks. A recognition trial was then completed after the delayed trial. The dependent variables were the number of words correctly recalled across the four learning trials, the immediate recall trial and the delayed recall trial, and the total number of correctly identified word on the recognition trial.

Intelligence. (i) Spot-the-Word task: This task is a computerized adaptation of the Spot the Real Word test (Baddeley et al., 1993). On each trial of this task, participants were presented with two words on the touch-screen. One of the two words was a valid word in the English language ("true" target word), and the second was a non-word foil. Participants were required to identify by touching the screen, which of the two words was the true target. The total correct score was the dependent measure.

Standard Neuropsychological Measures.

The standard tests were those developed previously, and described in detail in primary textbooks in the field of neuropsychology (see Lezak, 1995). These tests were selected according to two criteria: (1) the tests putatively measured the same cognitive construct as the tests of IntegNeuro, and (2) the tests were among the most common cognitive measures employed by clinical neuropsychologists (Lezak, 1995). Importantly, all dependent variables for the validity criterion measures were identified a priori. The standard neuropsychological measures were selected by two experienced clinical neuropsychologists based on the presumed overlap in tapping similar cognitive skills (convergent validity) and the lack of overlap in tapping the same cognitive skills (divergent validity).

The individual tests in equivalent domains are described in what follows.

Sensori-Motor Domain. (i) Finger Tapping: On this test participants were required to tap a counter with their dominant index finger as quickly as possible for 60 s. The dependent variable was the total number of taps recorded. Finger-tapping also served as the convergent validity variable for choice reaction time, as this response cannot be accurately captured by paper-and-pencil measures.

Attention Domain. (i) Spatial Span (Wechsler Memory Scale III; WMS-III; Wechsler, 1999b). Spatial Span is the visual analogue to the verbal digit span test. Participants observed the examiner touch a sequence of blocks, and they were required to touch the blocks in the same order. Both a forward and reverse order condition were administered. The dependent variable was the total number of correct trials. (ii) Digit Span test (WAIS-III; Wechsler, 1999a): Participants were administered the Digit Span subtest of the WAIS-III. Participants were presented listened to a sequence of digits (e.g., 4, 2, 7, etc.). The subject was then immediately asked to recall the digits. In the first part of the test, subjects were required to recall the digits in forward order (Digits forwards); in the second part, they were required to recall them in reverse order (Digits backwards). In each part, the number of digits in each sequence was gradually increased from 3 to 9, with two sequences at each level. The dependent measure for each part was the maximum number of digits the subject recalled without error. (iii) Adaptive Rate Continuous Performance Test (ARCPT; Cohen, 1993). This computerized test is based on the standard Continuous Performance Test, which is used to measure vigilance and sustained attention. A series of 1,000 letters are presented on the computer monitor, each for 100 ms. duration. Six out of every 100 letters consists of a target letter combination: “A” followed by “X” (A-X). Subjects respond by pressing the space bar. Accuracy is a function of hits, misses, discrimination ability (d’), and response bias (i.e., beta). Inter-stimulus interval (ISI) becomes shorter or longer on successive blocks of trials based on subjects’ accuracy, providing an information processing speed index. Trials are presented in 10 blocks of 100 trials, providing for measures of sustained attention. The dependent variable was the number of errors of omissions and false positives. (iv) Trail Making (Reitan, 1958): The Trail Making Test is a measure of combined visual search and cognitive flexibility. Trail Making A required participants to draw a line and connect 13 numbers (1–13), that are were scattered about the page, in ascending order. In the event that a subject committed an error (incorrect connection between numbers), the computer produced an audible signal and the subject was not allowed to connect additional numbers until the error was corrected. This process was explained thoroughly to the subject during the practice trial (which included an audible and visual example of an error). The dependent variable was time to completion.

Executive Function. (i) Trail Making (Reitan, 1958): Trail Making B required participants to connect 13 numbers (1–13) and 12 letters (A–L) in an alternating and ascending order. Time to completion was the dependent variable. (ii) Stroop (Golden, 1978): Participants were administered the Golden version of the standard Stroop test. Briefly, participants were required to name color words printed in incongruent ink as quickly as possible. The trial continued for 45 s. The dependent variable was total correct within the time period. (iii) Rey Complex Figure Test: The Rey Complex Figure Test (Meyers & Meyers, 1995) was also administered as a measure of visual learning and retention. Participants were instructed to copy the geometric figure as accurately as possible. Immediately after completing the copy trial, participants were asked to draw the figure from memory. Following a 20-min time delay, participants were again instructed to draw the figure from memory. No time limit was given and the drawings were scored according to standardized criteria (see Lezak, 1995). The dependent measure was the total score on the delayed trial. This dependent measure was selected to examine convergent validity on the Mazes task. While maze tasks are frequently identified as measures of pure executive function, the maze task in the IntegNeuro battery includes a memory component as well, because subjects are required to navigate the maze and retain the maze in memory for two successive trials without error. Due to this added memory component, Rey Delay were elected as the primary measure of convergent validity.

Language Domain. (i) FAS Fluency: Participants were required to generate words from the categories F, A, S as quickly as possible for 60 s on each trial. The dependent variable was total correct across the three letters. (ii) Animal fluency: Participants were required to name animals as quickly as possible for 60 s. The dependent variable was total correct.

Memory Domain. (i) California Verbal Learning Test-Research (CVLT-R; Delis et al., 1987): The CVLT-R was administered in standard fashion. Briefly, the test involves oral presentation of a 16-word list. The list items represent exemplars of one of four categories (fruit, clothing, tools, spices/herbs). Participants were read the list five times and asked to recall as many words as possible from the whole list on each of the trials. Participants were then read a distracter list, followed by free and cued recalls of the original list. After a 20-min delay, participants were asked to free recall the original list, followed by a cued recall trial, and then a recognition trial. The total number of words recalled across the five learning trials was the dependent measure for verbal learning. The dependent measures included total recall across the learning trials, recall on the short and long-delay free recall, and number of hits on the recognition trial.

Intelligence. (i) Wechsler Adult Intelligence Scale-III (Wechsler, 1999a). The WAIS-III was administered according to standard protocol. The entire battery of verbal and performance measures were administered to yield a full-scale Intelligence Quotient (IQ). The full scale IQ was computed as a reference to the estimated IQ derived from the Spot the Word test.

Data Analysis

The pairing of tests from the two batteries for purposes of validity testing is provided in Table 2. IntegNeuro™ data were scored using standardized and automated algorithms. Performances on the paper-and-pencil tests were scored according to established criteria. Evaluation of test performance on the paper-and-pencil tests was completed without any knowledge of the individual's performance on IntegNeuro™, in order to avoid any potential scoring bias.

Validity was assessed by examining the degree of similarity in performance on the IntegNeuro tests and the paper-and-pencil tests using correlational analyses, as well as predicted dissimilarity between the tests. Correlational analyses were computed for the entire group (50 individuals). Validity was also assessed by examining differences in performances on the individual tests between young individuals and older individuals. The influence of age, education, and sex was examined using correlation analyses and between-group contrasts, respectively.

RESULTS

Convergent and Divergent Validity

The correlational analyses are listed in Table 3 for the entire group of individuals. As evident in the Table, each test of IntegNeuro™ was significantly correlated with the paper-and-pencil measure previously developed to test the same cognitive construct. The correlation between IntegNeuro™ mea-

Table 2. Paired tests administered to assess the validity of IntegNeuro

IntegNeuro test	Convergent validity measure	Divergent validity measure
Digit Span total	Digit Span —WAIS III	Finger Tapping Nondominant
Visual Span total	Spatial Span—WMS III	CVLT-R Recognition
Finger tapping dominant	Finger Tapping Test—dominant	CVLT-R Long Delay
Finger tapping nondominant	Finger Tapping Test—nondominant	CVLT-R Long Delay
Memory recall on the learning trials	CVLT-R Learning Trials	Adaptive Rate Continuous Performance Test—Final ISI
Memory recall on the short delay	CVLT-R Short Delay	Rey Copy Trial
Memory recall on the delay trial	CVLT-R Long Delay Free Recall	Finger Tapping test dominant hand
Letter fluency	FAS	Finger Tapping nondominant
Animal fluency	Animal fluency	Digits Forward
Switching of attention A	Trail Making Test A	CVLT—Recognition
Switching of attention B	Trail Making Test B	Digit Span total
Mazes Time	Rey Complex Figure Test—delay	CVLT-R Short Delay
Verbal Interference	Stroop Incongruence trial	Rey Copy Trial
Choice Reaction Time	Finger Tapping Dominant*	Animal fluency
Working Memory Test	Adaptive Rate Continuous Performance Test	Spatial Span

*Finger tapping dominant hand was included as the convergent validity measure because standardized tests of choice reaction time are not commonly administered as part of paper-and-pencil neuropsychological batteries.

sures and at least one of the target convergent validity measures exceeded .53 in all cases, reflecting a statistically significant degree of overlap between the two variables. As evident in the Table 3, divergent validity was also supported as most of the IntegNeuro™ measures did not correlate with the divergent standard measures. The only exception was Switching of Attention B, which correlated with Digit Span. This is not a surprising finding given the impact of attentional processes on both tasks. Although this does not represent an exhaustive measure of the convergent and divergent validity parameters of the battery, the results provide some measure of specificity in terms of shared variance. Convergent validity was examined between Spot the Word and the Full Scale WAIS III IQ, and these two measures shared a significant amount of variance ($r = .76$). Divergent validity was not examined between the estimated intelligence measure and other cognitive functions as the former would be expected to correlate with most cognitive domains.

Table 3. Correlations between IntegNeuro measures and the paper-and-pencil tasks for all participants

IntegNeuro test	Convergent validity measure	Divergent validity measure
Digit Span	.53*	.20
Spatial Span	.63*	.27
Finger tapping dominant	.55*	.20
Finger tapping nondominant	.60*	.12
Memory recall on the learning trials	.67*	–-.26
Memory recall on the short delay	.76*	.14
Memory recall on the delay trial	.63*	.04
Letter fluency	.77*	.07
Animal fluency	.76*	.25
Switching of attention A	.53*	–.19
Switching of attention B	.65*	–.52
Mazes Time	–.65*	–.11
Verbal Interference	.70*	.22
Choice Reaction Time	–.53*	.24
Working Memory Test	.55*	–.00

*Statistically significant association between the two variables.

Relationships to Age, Sex, Education, and IQ

No significant differences were observed between males and females on any of the IntegNeuro™ subtests, although a trend was evident on nondominant finger tapping, with females recording slightly fewer taps compared to males (see Table 4). Correlational analyses were conducted to examine the relationships between age, education, and performance on the individual measures of IntegNeuro™. Results from these analyses revealed that only digit span correlated with number of years of education ($r = .44$), with higher scores associated with more years of education. More consistent relationships were evident between performance on the IntegNeuro™ tests age and IQ. Age correlated significantly with reverse digit span forward ($r = –.56$), digit span backward ($r = –.44$), dominant tapping ($r = –.34$), nondominant tapping ($r = –.31$), choice reaction time ($r = .31$), maze time to completion ($r = .54$), animal fluency ($r = –.61$), and FAS fluency ($r = –.39$).

Full Scale IQ as determined by the WAIS-III was significantly correlated with performance on the IntegNeuro™ digit span test ($r = .51$), choice reaction time ($r = .29$), verbal interference test ($r = .47$), total verbal learning

Table 4. Performance differences between males and females

	Males		Females		
Measure	Mean	(SD)	Mean	(SD)	*F*
Digit Span	11.0	(2.4)	10.0	(3.1)	2.1
Visual Span	7.1	(2.5)	7.2	(2.3)	.04
Finger tapping dominant	157.7	(24.3)	156.7	(33.3)	.01
Finger tapping nondominant	145.41	(18.5)	156.8	(20.2)	3.6 p = .06
Learning Trials	29.7	(7.2)	31.6	(5.7)	1.0
Learning Short Delay	6.4	(2.7)	7.2	(2.3)	.28
Learning Long Delay	5.9	(2.8)	6.8	(2.5)	1.2
Letter fluency	47.5	(12.1)	47.5	(13.2)	0.0
Animal fluency	22.9	(6.1)	23.0	(5.6)	0.0
Switching of attention A	24.2	(8.1)	23.4	(8.1)	0.0
Switching of attention B	45.6	(12.1)	46.6	(12.9)	0.0
Mazes time	27.3	(8.0)	28.4	(13.6)	0.0
Verbal interference	19.7	(5.1)	19.9	(4.4)	0.0
Choice reaction time (ms)	757.2	(161.5)	743.3	(147.2)	0.0
Working memory	1.6	(1.1)	2.8	(2.4)	2.7

(r = .50), long delayed verbal recall (r = .39), Switching of Attention 2 (–.42), Mazes time to completion (r = –.46), Animal fluency (r = .43), and letter fluency (r = .63).

DISCUSSION

The results of the study provide preliminary support for the use of IntegNeuro™ to assess cognitive function. The correlational analyses revealed strong relationships between the tests of IntegNeuro™ and standard measures of cognitive function. In addition, divergent validity was demonstrated by the absence of significant relationships between the target measures and cognitive tests believed to tap different underlying constructs on most of the tests. It is important to note that complete divergence is difficult to attain given the dominance of attention and speed factors inherent in the test battery, and the association of these skills with most higher-order cognitive tasks. Nevertheless, the findings provide preliminary support for the convergent and divergent validity of the computerized battery.

It is worth noting that the strongest correlations were observed for the memory measures and the fluency measures, which required subjects to hear and comprehend accurately verbal instructions and memory lists, as well as provide verbal responses that were subsequently recorded via the attached microphone. A concern regarding these procedures is whether voice activation software is capable of accurately detecting the responses. Although not directly examined in the current study, the strong correlations between verbal fluency and memory scores across the computerized and noncomputerized measures suggest that this is not a significant problem. Overall the findings support the use of language and verbal memory measures in computerized assessment.

IntegNeuro™ may provide clinicians with a sensitive screening of core cognitive abilities that are vulnerable to degenerative disease and other forms of brain injury. Importantly, the battery includes tests that are sensitive to deficits associated with Alzheimer's Disease (AD; long-term recall and recognition memory; Tierney et al., 2001; Bondi & Monsch, 1998), vascular dementia (fluency and recognition memory; Tierney et al., 2001; Collie & Maruff, 2000), and conversion from mild cognitive impairment to probable AD (Trail Making B and verbal learning; Albert et al., 2001; Small et al., 1997). Tests are also included that differentiate AD from probable vascular cognitive impairment (letter fluency and recognition memory; Tierney et al., 2001). As such, the battery is potentially capable of detecting clinically meaningful declines in cognitive function, although this needs to be determined empirically in future studies.

The strength of the correlations appears between the individual tests of IntegNeuro™ and the standard measures appears consistent with the correlations reported between other computerized measures and indices of convergent validity. For example, correlations between measures of the computerized NES 3 and standard paper and pencil measures ranged form .20 to .70 (Proctor et al., 2000), whereas the measures of IntegNeuro™ ranged from .53 to .77. IntegNeuro™ may offer some advantages compared to other computerized batteries concurrently available. Specifically, the inclusion of language-based measures including verbal fluency and verbal memory provide opportunities to assess key cognitive domains most consistent with routine clinical practice, as well as examine cognitive domains that are vulnerable to early degenerative decline. Further, the similar task demands between the computerized tests and the standard measures (e.g., Switching of Attention and Trail Making) allow for a more direct comparison of individual skills.

One area of future research regarding this and all computerized batteries is the issue of computer familiarity. Computer familiarity may account for as much as 40% of the variance in performance on computerized cognitive batteries (Browndyke et al., 2003; Weber et al., 2002). This effect may be more of an issue for older individuals, who have less personal history working with computers. The methodology incorporated into IntegNeuro™ may reduce the confounding influence of computer familiarity because the apparatus does not require a keyboard or mouse. However, this remains conjecture until studied directly in future investigations.

It is worth noting that participants in the current study reported a high level of education, and this may have influenced the overall findings. Specifically, the higher education values of the sample provide warrants caution in the generalization of the findings to individuals with lower levels of education. Further, it is possible that the restricted range of the education values in this sample constrained the correlation coefficients between education and performance on the cognitive measures. Future studies with less highly educated participants will be an important step in the subsequent validation of the battery. It will also be important to examine relationships between individual tests of IntegNeuro™ and performance on additional "traditional" measures of cognitive function. For example, there are a number of indices that can be derived from the vigilance measure of IntegNeuro™ and the ARCPT (Cohen, 1993), and it would be interesting to examine the relationships between the verbal memory measure and performance on tests of prose passages, as well as other list learning tasks with similar learning and retention requirements (e.g., from the WMS-III, Wechsler, 1999b).

In summary, this preliminary investigation of the validity of IntegNeuro™ provides encouraging evidence for the utility of this measure. The battery appears to capture essential elements of standardized cognitive assessments including language-based measures. The standardization of task procedures and instructions should offer a significant benefit to multi-site studies and research initiatives. Additional information regarding the reliability of the battery, the utility of alternate forms of the tests, the sensitivity of the tests to clinical indications, and differences in education and computer familiarity will be important next steps to determine the broader utility of the battery for research and clinic applications.

REFERENCES

Albert, M. S., Moss, M. B., Tanzi, R., & Jones, K. (2001). Preclinical prediction of AD using neuropsychological tests. *Journal of the International Neuropsychological Society*, *7,* 631–639.

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual for Mental Disorders, Fourth Edition.* Washington, DC: Author.

American Psychological Association. (1987). Guidelines for computer based tests and interpretations. Washington, DC.

Baddeley, A., Emslie, H., & Nimmo-Smith, I. (1993). The Spot-the-Word test: A robust estimate of verbal intelligence based on lexical decision. *British Journal of Clinical Psychology*, *32,* 55–65.

Baker, E. L., Letz, R. E., Fidler, A.T., Shalat, S., Plantamura, D., & Lyndon, M. A. (1985). Computer-based neurobehavioral evaluation system for occupational and environmental epidemiology: Methodology and validation studies. *Neurobehavioral Toxicology and Teratology, 7,* 369–377.

Bondi, M. W., & Monsch, A. U. (1998). The role of memory assessment in preclinical Alzheimer's disease. In A. I. Troster (Ed.), *Memory in neurodegenerative disease: Biological, cognitive and clinical perspectives* (pp. 278–289). Cambridge University Press.

Browndyke, J. N., Albert, A. L., Malone, W., Schatz, P., Paul, R. H., Cohen, R. A., Tucker, K. A., & Gouvier, W. D. (2002). Computer-related anxiety: Examining the impact of technology-specific affect on the performance of a computerized neuropsychological assessment meaure. *Applied Neuropsychology, 9,* 210–218.

Cohen, R. (1993). *Neuropsychology of attention*. New York: Plenum Press.

Collie, A., & Maruff, P. (2000). The neuropsychology of preclinical Alzheimer's disease and mild cognitive impairment. *Neuroscience and Biobehavioral Reviews*, *24,* 365–374.

Delis, D., Kramer, J., Kaplan, E., & Ober, B. (1987). *Manual: California Verbal Learning Test, Adult Version.* San Antonio: The Psychological Corporation.

Devivo, K., Rothland, J., Price, L., & Fein, G. (1999). Computerized assessment of arithmetic computation skills with MicroCog. *Journal of the International Neuropsychological Society, 3,* 199–200.

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

Golden, C. J. (1978). *Stroop Color and Word Test*. Chicago: Stoelting Company.

Gordon, E., Cooper, C., Rennie, D., & Williams, L. M. (2005). Integrative neuroscience: The role of a standardized database. *Clinical EEG and Neuroscience, 36,* 64–75.

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

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

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research & Therapy*, *33*(3), 335–343.

Meyers, J. E., & Meyers, K. R. (1995). *Rey Complex Figure Test and Recognition Trial. Professional Manual.* Psychological Assessment Resources, Lutz, Florida.

Morris, R. G., Evenden, J. L., Sahakian, B. J., & Robbins, T. W. (1986). 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 & E. C. Goodman (Eds.), *Cognitive neurochemistry* (pp. 21–36). Oxford: Oxford University Press.

Osterrieth, P. A. (1944). Le test de copie d'une figure complex: Contribution a l'etude de la perception et de la memoire. *Archives of Psychologie*, *30,* 286–356.

Proctor, S. P., Letz, R., & White, R. (2000). Validity of a computer-assisted neurobehavioral test battery in toxicant encephalopathy. *Neurotoxicology, 21*(5), 703–714.

Reitan, R. M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and Motor Skills*, *8,* 271–276.

Small, B. J., Herlitz, A., Fratiglioni, L., Almkvist, O., & Backman, L. (1997). Cognitive predictors of incident Alzheimer's disease: A prospective longitudinal study. *Neuropsychology*, *11,* 413–420.

Spreen, O., & Strauss, E. (1988). *A compendium of neuropsychological tests.* New York: Oxford University Press.

Tierney, M. C., Black, S. E., Szalai, J. P., Snow, W. G., Fisher, R. H., Nadon, G., & Chui, H. C. (2001). Recognition memory and verbal fluency differentiate probable Alzheimer disease from subcortical ischemic vascular dementia. *Archives of Neurology*, *58,* 1654–1659.

Walsh, K. W. (1985). *Understanding brain damage: A primer on neuropsychological evaluation* (pp. 234–237). Edinborough: Churchill Livingstone.

Weber, B., Fritze, J., Schneider, B., Kuhner, T., & Maurer, K. (2002). Bias in computerized neuropsychological assessment of depressive disorders caused by computer attitude. *Acta Psychiatrica Scandinavica*, *105,* 126–130.

Wechsler, D. (1999a). *Wechsler Adult Intelligence Scale-Three*. San Antonio: Psychological Corporation.

Wechsler, D. (1999b). *Wechsler Memory Scale-Three*. San Antonio: Psychological Corporation.

Williams, L., Simms, E., Clark, C., Paul, R., Rowe, D., & Gordon, E. (Submitted). The reproducibility of a standardized and integrated neurophysiological and neuropsychological test battery. *International Journal of Neuroscience.*

Wittchen, H. U. (1991). Reliability and validity studies of the WHO Composite International Diagnostic Interview (CIDI): A critical review. *Journal of Psychiatric Research*, *28*, 57–84.

Wittchen, H. U., Robins, L. N., Cottler, L. B., Sartorius, N. et al. (1999). Cross-cultural feasibility, reliability and sources of variance of the Composite International Diagnostic Interview (CIDI). *British Journal of Psychiatry*, *159,* 645–653.