Archives of Clinical Neuropsychology 21 (2006) 449–467

Archives of CLINICAL NEUROPSYCHOLOGY

Standardized assessment of cognitive functioning during development and aging using an automated touchscreen battery

C. Richard Clark [a,*], Robert H. Paul [b], Leanne M. Williams [f,g,h], Martijn Arns [c,d], Kamran Fallahpour [c,d,e], Carolyn Handmer [c], Evian Gordon [c,f,g]

[a] School of Psychology, Flinders University, P.O. Box 2100, Adelaide, SA 5001, Australia
[b] Brown Medical School, Department of Psychiatry, Centers for Behavioral and Preventive Medicine, United States
[c] The Brain Resource International Database, Brain Resource Company, Australia
[d] BraInquiry, Nijmegen, Netherlands
[e] Institute of Neuropsychology and Cognitive Performance, NY, United States
[f] Department of Psychological Medicine, University of Sydney, Australia
[g] The Brain Dynamics Centre, Westmead Hospital, Australia
[h] School of Psychology, University of Sydney, Australia

Accepted 16 June 2006

Abstract

This study examined the effects of age, gender and education on subjects spanning nine decades on a new cognitive battery of 12 tests. One thousand and seven participants between 6 and 82 completed the battery under standardized conditions using an automated, computerized touchscreen. Sensitive indicators of change were obtained on measures of attention and working memory, learning and memory retrieval, and language, visuospatial function, sensori-motor and executive function. Improvement tended to occur through to the third and fourth decade of life, followed by gradual decrement and/or stabilized performance thereafter. Gender differences were obtained on measures of sustained attention, verbal learning and memory, visuospatial processing and dexterity. Years of education in adults was reflected in performance on measures of verbal function. Overall, the test battery provided sensitive indicators on a range of cognitive functions suitable for the assessment of abnormal cognition, the evaluation of treatment effects and for longitudinal case management.

Keywords: Neuropsychology; Development; Ageing; Education; Gender; Cognition

Neurocognitive assessment has continued to serve as a primary method to study both the development and the degeneration of the brain, and improvements in the testing process have helped to solidify its role in the field of neuropsychology. The introduction of standardized methods of test administration, scoring and interpretation represents one of the most important advances in the past century. At present, scientists and clinicians have a remarkable range of standardized tests from which to choose, but individual tests are often constrained by a lack of adequate normative data or limited coverage of domains of function, leaving significant opportunities for further improvement.

This paper reports the development of a cognitive test battery designed to permit measurement across the full lifespan (Gordon, 2003a, 2003b). It has been well reported that cognitive abilities increase with age up until the third or

* Corresponding author. Tel.: +61 8 8201 2425; fax: +61 8 8201 3877.
E-mail address: Richard.clark@flinders.edu.au (C.R. Clark).

0887-6177/$ – see front matter
doi:10.1016/j.acn.2006.06.005

fourth decade of life (e.g. Anderson, Anderson, Northam, Jacobs, & Catroppa, 2001; Arceneaux, Hill, Chamberlin, & Dean, 1997; Korkman, Kemp, & Kirk, 2002) whereupon they plateau and then gradually decline (e.g. Brabeck, 1984; Compton, Avet-Compton, Bachman, & Brand, 2003; Costa & McCrae, 1993; Herb, 1995; Rabbitt, 2002; Salthouse, 2000, 2001). It is also reported that females develop more quickly in some areas and that, broadly speaking, males develop to perform better on motor tasks and females on some verbal and memory tasks. Further, the relative levels of cognitive performance generally increase with higher levels of education.

Individual studies addressing longitudinal change in cognitive ability tend to focus either on the developmental years or on the effects of aging but not both. The present study used a standardized, computerized test battery with reported reliability and validity (Paul et al., 2005; Williams et al., 2005) to examine healthy participants between the ages of 6 and 82. This battery taps the primary cognitive domains including attention, learning and memory, speed of processing, verbal versus visuomotor ability and executive function. A key aim was to obtain a broad profiling of cognitive ability that considered cross-sectional effects of the full age range within a single study across and to differentiate effects according to gender and years of education. Thus, this study employed the same broad-based, cognitive battery on all participants across the full age range. In particular, the study permitted concurrent assessment of variation in the peaks of the development trajectory across different cognitive domains, the examination of gender differences along these trajectories and examination of the relationship between observed development patterns over time on the various cognitive domains and the subsequent effects of ageing on those domains.

Three methodological issues were also addressed in this study. Firstly, the limitation of small normative samples drawn from a single demographic was addressed by administering the test battery at a number of laboratories to over 1000 individuals. Secondly, the testing environment was tightly standardized. Standardization included the physical environment within which participants are tested (lighting, ambient temperature, visual and auditory distractions), the computerized basis of the control of test delivery and for the measurement of performance and the automation of test instructions. The provision of such controls limited the impact of the environment and human interaction on participant performance. Thirdly, the test battery was constructed to be completed in under an hour, with test order kept invariant. These two controls limited the effects of factors, such as vigilance, motivation and fatigue on test performance, whilst ensuring that the impact of such factors is standardized across tests.

The computerized tests forming the battery were based on existing and well recognized paper and pencil tests known to be sensitive to brain dysfunction. Computerized assessment has increasingly received recognition for use in the research and clinical setting. 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. 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.

The extent of standardization in the design of the battery used in the present study, together with the relatively large size of the normative database, potentially offers good sensitivity to the measurement of neurocognitive function. This paper specifically addresses the sensitivity of the battery to the effects of age, gender and level of education on cognitive abilities for each cognitive domain across the full age span. A quadratic effect of age was generally expected to reflect improving capabilities over the years of early development followed by decline during old age. No general hypothesis was levied with respect to gender and level of education, with the effects of these variables expected to vary across the measures obtained. The reliability and validity of the battery (published as IntegNeuro: see www.brainresource.com) has been reported recently.

1. Method

1.1. Participants

Participants were 1007 male and female volunteers, spanning nine decades (6–82 years; mean age 28.7). Six laboratories from Australia (Adelaide, Sydney), USA (NY, Rhode Island) and Europe (UK, Holland) participated in the data acquisition in a standardized manner with identical hardware, software, paradigms and experimental procedures. The data from the participants reported in this paper form part of a work in progress related to formation of an international database of brain and cognitive function (Brain Resource International Database) with final numbers to be spread evenly across age range and gender. A recent comparison of data obtained to date on the battery from the laboratories cited above failed to identify significant differences in cognitive function across the three continents (Paul et al., 2006). All participants completed a screening questionnaire to provide demographic information as well as a self-report of psychological, neurological and physical history. The completion of the questionnaire was overseen by a parent/guardian in the case of participants less than 18 years of age, but with assistance offered to such participants only in relation to specific questions.

Exclusion criteria included a personal history of mental illness, physical brain injury, neurological disorder or other serious medical condition and/or a personal history of sleep disorder, learning disability, or drug or alcohol addiction. Subjects were also excluded if they had a family history of Attention Deficit Hyperactivity Disorder (ADHD), schizophrenia, bipolar disorder or genetic disorder. The SPHERE questionnaire (Hickie, 1998) was used to screen out individuals with a likely anxiety or depressive disorder. Subjects were required to refrain from caffeine-intake and from smoking for at least 2 h prior to testing to standardize the effects of these variables. All subjects (or their guardians for subjects less than 18 years of age) provided written informed consent to participate in the database.

1.2. Materials and equipment

Participants were seated in front of a touchscreen computer (NEC MultiSync LCD 1530 V) located within a sound and light attenuated room with an ambient temperature of approximately 24 °C. The touchscreen was located centrally on a desktop in front of participants with the touch surface subtending an angle of 15°. Task instructions and materials were pre-recorded and delivered in a standardized way using computer 'wav' files presented via headphones and using the visual display of a touchscreen computer. An iterative, automated protocol was used to ensure task comprehension and compliance. This involved a computerized protocol directing the participant through a number of practice trials on each test. If three practice trials were failed on a test, then the participant was taken automatically to the next test in the battery. Task instructions are provided concurrently in verbal form (by headphones) and in written form (on the touchscreen) to increase the accessibility of the test battery to the younger participants in the study, where reading ability may be less well formed.

1.2.1. Integneuro test battery

The test battery consists of 12 tasks that take approximately 50 min to complete. Instruction and practice is completed immediately prior to attempting the task. The tests cover five cognitive domains as follows.

1.2.1.1. Sensori-motor. The Motor Tapping test is a variation of the Finger Tapping test that is considered to provide an index of manual dexterity (Halstead, 1947; Lezak, 2004). The participant is required to place the palm of the hand on the touchscreen and tap with the index finger as fast and as often as possible for a period of 30 s. Measures of tapping frequency from the dominant hand are reported.

The Choice Reaction Time test: four black circles are presented equally spaced along a semicircular arc across the top of the touchscreen, with two circles located to the left of midline and the other two on the other side of midline. The participant is required to rest the pointing finger of the hand with which they write on a white circle positioned on the midline of the screen and located below and equidistant from each of the black circles. Periodically and with equal probability, a black circle is illuminated in green. The participant is required to lift their finger and touch each illuminated circle as fast as they can. A measure of the average speed of response is obtained.

1.2.1.2. Learning and memory. The Memory Recall and Recognition test is a variant of the Rey Auditory Verbal Learning and Memory task (e.g. Geffen, Moar, Hanlon, Clark, & Geffen, 1990; Rey, 1964) commonly used to provide measures of auditory-verbal learning, recall and recognition, as well as indices of self-monitoring ability. The participant is presented with a sequence of 12 words binaurally via headphones, one second at a time (Learning Trial 1). Prior to the reading of the list, the participant is instructed to say back, immediately after the list has been read, as many of the words as they can remember from the list, in any order. This procedure is then repeated three more times, with the same instructions (Learning Trials 2–4). The total number of words recalled during these four trials (Trials 1–4) is recorded. The participant is then presented with a second list of words and asked to say back as many of these as can be remembered (Distractor Trial). None of the words in the second list are phonetically or semantically related to the words in the first list. Immediately following this trial, the participant is asked to recall as many words as can be remembered from the first list (Immediate Recall Trial). The task takes about 6 min to this point. Approximately 25 min later, after completing a number of other tests in the battery, the participant is again asked to recall as many words as possible from the first list (Delayed Recall Trial). The subject is then presented one at a time with a series of 24 words on the computer screen (Recognition Trial). Half of these words are the words from the first list; the remaining words are new words. The words are in fixed, pseudorandom order. Following each word, the participant is required to touch a "Yes" or "No" button on the touchscreen according to whether or not the word was in the first list.

All words are concrete words between four and seven letters in length, drawn from the English language. Concreteness was determined from the merged norms of Paivio (Paivio, Yuille, & Madigan, 1968), Colerado (Toglia & Battig, 1978) and Gilhooly–Logie (Gilhooly & Logie, 1980). Using this derivation, only words with a concreteness rating of 538 or more were selected (Paivio et al., 1968). The value of 538 represents the mean concreteness value of the derived list plus one standard deviation. The words in each list are of high frequency, with a value of 50 or more in written frequency (Kucera & Francis, 1967). There are no semantic or phonemic similarities between words either within or across lists. All lists are closely matched on mean concreteness, mean number of letters and mean frequency. Subject verbal recall is recorded in 'wav' files.

The Maze task is a computerized adaptation of the Austin Maze (Walsh, 1985) used to assess high level mental functions, such as planning, foresight and self-monitoring during the course of learning and remembering a complex maze. A computerized format for this task has been shown to be equally as effective as the conventional task design using a light-up circuit board (Morrison & Gates, 1988). An 8×8 grid of red circles (the maze) is displayed across the computer screen in a square array. The maze contains a hidden path commencing at a marked circle (in yellow) on one of its edges and finishes at a marked circle (in blue) on another of its edges. The participant's task is to discover and remember this path. The participant is instructed to move through the maze by touching the up, down, left and right arrows displayed in congruent pattern immediately below it. Each learning trial begins at the yellow start point and is completed once the blue finish button is reached. Participants are required to perform learning trials until they either complete the maze twice in a row without error or 7 min have passed (which ever comes first).

This adaptation of the Austin Maze introduces a number of variations. Firstly, it is not possible for the participant to move diagonally within the maze. Secondly, the maze is smaller, with each side containing 8 rather than 12 grid points. Finally, the duration of the task is limited to 7 min; the original version allowed the participant to perform the task for an indefinite period of time.

This adaptation of the task yields the following measures relevant to visual spatial learning and memory: the number of trials completed or until time out occurred (Trials completed) and the time to task finish (Time to Complete).

1.2.1.3. Language. The Letter Fluency test is a variant of the Controlled Oral Word Association (COWA) test, or FAS, as it is also known (Benton & Hamsher, 1989). The test is used to provide an index of verbal fluency, as measured by the quantity of words produced within given letter categories. This task requires the participant to recall as many words as possible beginning with each of the letters F, A and S, in that order. One minute is allowed for each letter. Subject verbal recall is recorded in 'wav' files. The dependent measure is the number of words recalled across the three letters.

The Spot the Real Word task is a computerized adaptation of the Spot the Word test (Baddeley, Emslie, & Nimmo-Smith, 1993). In each of a series of trials, a pair of words is presented side by side on the computer screen. One of the words in each pair is a real word; the other word is a pseudo word, with the ordering of words pseudo randomized over trials. The words in each pair remain on the screen until the subject indicates by touch which of them is considered to be the real word. The dependent variable reported in this study is the number of words correctly selected. Note the total correct score can also be entered into a regression formula that factors education and age to render an estimated intelligence quotient for those 16 years and over (Sullivan, Senior, & Hennessy, 2000).

1.2.1.4. Attention and working memory. The Span of Visual Memory task is adapted from the Corsi Blocks task (Milner, 1970) and another of its variants, the Dot Location task (Roth & Crosson, 1985). The Corsi Blocks test is used to provide an index of visual short-term memory capacity (Lezak, 2004). The adaptation addresses a noted difficulty with previous versions of the test that confound sequence length with path length. In the present task, path length increases with sequence length. Nine asymmetrically positioned squares are displayed constantly on the computer screen for the duration of this task. The task consists of a series of trials in which a number of the squares flash briefly in sequence, followed immediately by the sound of a tone. The participant is required to touch the squares that flashed in sequence order, with only one attempt per trial. Sequence length varies between two and nine, with two trials for each length and with trials presented in ascending sequence order. The task is terminated either when the participant fails 2 trials of any sequence length or when all 18 trials are completed. The dependent measure is the longest sequence length correctly completed.

Part 1 of the Switching of Attention (SOA) test is part of a computerized adaptation of the Trail Making test (Reitan, 1955, 1958). Part 1 provides an assessment of visuomotor tracking and motor speed, in which the subject is presented with a display of 25 numbers on the computer screen and asked to touch the numbers in ascending numerical sequence (i.e. 1 2 3 ...). The numbers are drawn without replacement from the range 1 to 25 and are displayed in a fixed pseudo random pattern. The dependant variable reported is the time taken to complete the test successfully (Time to Completion).

The Time Estimation task consists of a series of trials in which a black circled displayed on the computer screen turns green for a number of seconds. For each trial, the participant is required to indicate the duration in seconds of the colored stimulus. Response is carried out using a fixed display touchpad showing the numbers 1–12 in sequence from left to right across the computer screen. Stimulus duration ranges equiprobably between 1 and 12 s, with order of presentation being pseudo random over trials. Proportional bias in time estimation is estimated from the absolute value of the average difference between the actual duration of the stimulus and the users estimate weighted by the length of the stimulus.

To tap sustained attention, a series of letters (drawn equiprobably from the set of B, C, D and G) were presented to the subject on the computer screen (for 200 ms), separated by an interval of 2.5 s (Sustained Attention task). 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 reported are reaction time (Reaction Time) and number of false positive responses (False Positives).

The Digit Span task is an adaptation of earlier tests of immediate verbal working memory (Lezak, 2004). It consists of two sub-tests. Firstly, a test of Forward Digit Span, which has been used to provide a measure of immediate memory recall. The second is a test of Reverse Digit Span, which has been used to provide a working memory index of mental tracking ability (Lezak, 2004). The Forward Digit Span task consists of a number of trials in which a series of numerals are presented at a constant rate in a box on the computer screen. Immediately after each trial, the participant is required to touch the numerals on a screen keypad, which then appears on the screen, in the order in which they were flashed. There is a 5 s delay between trials. Sequence length varies between three and nine, with two trials for each length and with trials presented in ascending sequence order. The task is terminated either when the participant fails 2 trials of any sequence length or when all 18 trials are completed. The score is the longest sequence length correctly completed. The Reverse Digit Span task is identical to the Forward Digit Span task, except that numbers must be recalled in the reverse order of presentation. The dependent measures reported are the longest sequence length correctly completed in each case (Forward Digit Span score; Reverse Digit Span score).

Part 1 of the Word Interference test is part of an adaptation of the Stroop test (Stroop, 1935). This task consists of a series of colored words presented one at a time on the computer screen. Below the color words are the four words red, yellow, green and blue, displayed permanently in black across the computer screen. The color and name of each colored word is always one of these four colors with the constraints that no colored word has the same color and name but that each name and color is equiprobable. Colored words are presented in pseudo random order. In Part 1 of this test, the participant is required to identify the name of each colored word as quickly as possible after it is presented on the screen. Colored words remain on the screen until the subject responds. This part lasts for 1 min. The number of words correctly identified (Score) is recorded.

1.2.1.5. Executive function/planning. In Part 2 of the Switching of Attention task, the subject is presented with a pattern of 13 numbers (1–13) and 12 letters (A–L) on the screen and is 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 reported is the time to completion (Time Part 2).

Part 2 of the Word Interference test measures the ability to suppress automatic, well-learned responses in the face of competing demands. In this part, the participant is required to name the color of each colored word as quickly as possible after it is presented on the computer screen. Colored words remain on the screen until the subject responds.

This part lasts also lasts for 1 m. The dependent measure obtained is the number of words correctly identified (Score Part 2). It should be noted that this version of the Stroop test differs qualitatively from the original in terms of the nature of response. Instead of verbalizing the name/color, as in the original versions, the subject selects a name/color from a four choice visual array placed below the probe color word on the screen. In principle, this could allow the subject to choose by visual match rather than engaging in mental search before articulation. Nevertheless, the process of suppression of automatic response would still be required. In addition, the task has a fixed duration. This means that absolute and/or relative task performance is not measured in terms of the time to complete a fixed number of trials.

Two measures of executive functioning are also obtained from the Maze task. These include the total number of off-path moves (errors) and the total number of off-path moves made the time a path turn should have been made (total overruns). Two measures of executive functioning obtained from the Verbal Memory Recall task are the number of words incorrectly recalled during trials 1–4 (Intrusions) and the number of repeats of correctly recalled trials during trials 1–4 (Repeats). These were taken as indices of self-monitoring ability.

1.3. Data analysis

The 12 tests yielded 47 dependent variables for analysis, of which 24 are reported in this paper. Scoring of most of these variables is automated. Manual scoring was required for the two tasks involving verbal response (Memory Recall; Word Generation). There were a number of missing values for some participants; however, these data were not replaced.

Data were divided into 10 age groups (age group) as follows: 6–8, 9–11, 12–14, 15–19, 20–29, 30–39, 40–49, 50–59, 60–69 and over 70. These divisions were asymmetrical over the age range to reflect the relatively rapid rates of development during the early years. Data were then analysed in SPSS using a full factorial general linear model (GLM) to investigate the effects on performance measures of the independent variables of age group, gender and years of education.

1.3.1. Analysis of age group and gender

The distribution of participants over the 10 age groupings was as follows: 6–8 years ($n = 57$; female (F): 28, male (M): 29; mean years of education (YOE): 2.7), 9–11 ($n = 98$; F: 46, M: 52; YOE: 5.5), 12–14 ($n = 83$; F: 34, M: 49: YOE: 8.4), 15–19 ($n = 146$; F: 63, M: 83; YOE: 12.3), 20–29 ($n = 254$; F: 130, M: 124; YOE: 13.9), 30–39 ($n = 114$; F: 48; M: 66; YOE: 14.1), 40–49 ($n = 97$; F: 59, M: 38; YOE: 13.4), 50–59 ($n = 83$; F: 49, M: 34; YOE: 13.3), 60–69 ($n = 44$; F: 21, M: 23; YOE: 12.2), >70 ($n = 31$; F: 15, M: 16; YOE: 12.0). Overall, the sample comprised 514 males and 493 females. A GLM was undertaken with age group and gender as fixed factors, with a polynomial contrast for age group and a simple contrast for gender.

1.3.2. Analysis of years of education

Only participants 20 years of age or more were included in this analysis to avoid the confound between years of education and age in younger participants due to the positive correlation of these variables. The participants were assigned to one of four groups representing general level of education: early school leavers (<10 years of education; $n = 82$), high school (10–12 years; $n = 130$), tertiary (13–15; $n = 157$) and postgraduate (>15 years; $n = 254$) level of education.

2. Results

2.1. Age group

Significant linear and/or quadratic main effects of age were obtained on most measures (see Table 1). The only variable that did not differ across age group was a measure of executive function (repeated recalls during the Memory Recall task).

2.1.1. Attention and working memory

All measures of attention and working memory were fitted by a quadratic trend over age group. Figure 1 shows, with minor exception, a graded improvement in performance over all measures through to the middle age bands (15–39) after

Table 1
Statistical effects of age group, gender and years of education on tests of attention and working memory, executive function, language, learning and memory and sensori-motor function

	Age group			Gender	Age × Gender	Education	
	F	Linear (S.E.)	Quadratic (S.E.)	*F*	*F*	*F*	Linear (S.E.)
Attention and working memory							
Time estimation—Bias	3.23***	0.521**	0.513*	1.27	0.91	0.404	NS
Switching of Attention—Time Part 1	7.95***	NS	2508***	2.47	1.13	2.39	NS
Sustained Attention—False Positives	4.29***	1.04***	1.03***	11.42***	1.52	0.135	NS
Sustained Attention—Reaction Time	15.01***	NS	64.21***	28.29***	12.7***	3.08**	NS
Visual Memory Span—Span	4.5***	NS	0.621***	0.024	0.698	1.65	NS
Reverse Digit Span—Score	4.56***	NS	0.802***	0.063	0.948	5.82***	0.157*
Verbal Interference—Score Part 1	6.44***	1.21*	1.19***	0.385	0.26	2.4	NS
Forward Digit Span—Score	3.19***	NS	0.741**	0.784	0.412	5.23***	0.124**
Executive function							
Switching of Attention—Time Part 2	11.97***	NS	4458***	0.054	0.365	3.25*	882.9**
Maze—Errors	14.31***	14.82***	14.6***	30.34***	8.04***	1.66	NS
Maze—Overruns	13.95***	5.79***	5.7***	28.43***	8.85***	2.18	NS
Verbal Interference—Score Part 2	13.05***	1.56***	1.54***	0.004	0.695	0.948	NS
Memory Recall and Recognition—Repeats	1.59	NS	NS	4.76*	1.88	3.15*	1.39*
Memory Recall and Recognition—Intrusions	1.62	NS	NS	1.52	0.53	0.197	NS
Language							
Word Generation—Score	35.74***	2.97***	2.91***	0.064	1.65	5.04**	1.37**
Spot the Real Word—Score	67.86***	0.76***	0.719***	0.001	0.48	8.8***	0.632***
Learning and memory							
Memory Recall and Recognition—Trials 1–4	11.13***	NS	1.28***	6.96**	0.682	1.25	NS
Memory Recall and Recognition—Recognition	7.46***	NS	0.229***	2.23	1.52	1.35	NS
Memory Recall and Recognition—Immediate Recall	10.38***	NS	0.479***	8.72**	0.703	0.386	NS
Memory Recall and Recognition—Delayed Recall	10.68***	0.467*	0.447***	10.41***	1.52	0.205	NS
Maze—Time to Complete	11.66***	18948***	18030***	0.054	1.04	2.74*	1558.47**
Maze—Trials Completed	3.76***	NS	1.89*	0.982	1.8	0.28	NS
Sensori-motor							
Motor Tapping—Taps (dominant)	33.43***	3.68***	3.5***	22.28***	1.25	1.17	NS
Choice Reaction Time—Mean Reaction Time	2.96**	NS	121.04*	1.49	1.74	2.53	NS

NS, not significant.
* $p \leq 0.05$.
** $p \leq 0.01$.
*** $p \leq 0.001$.

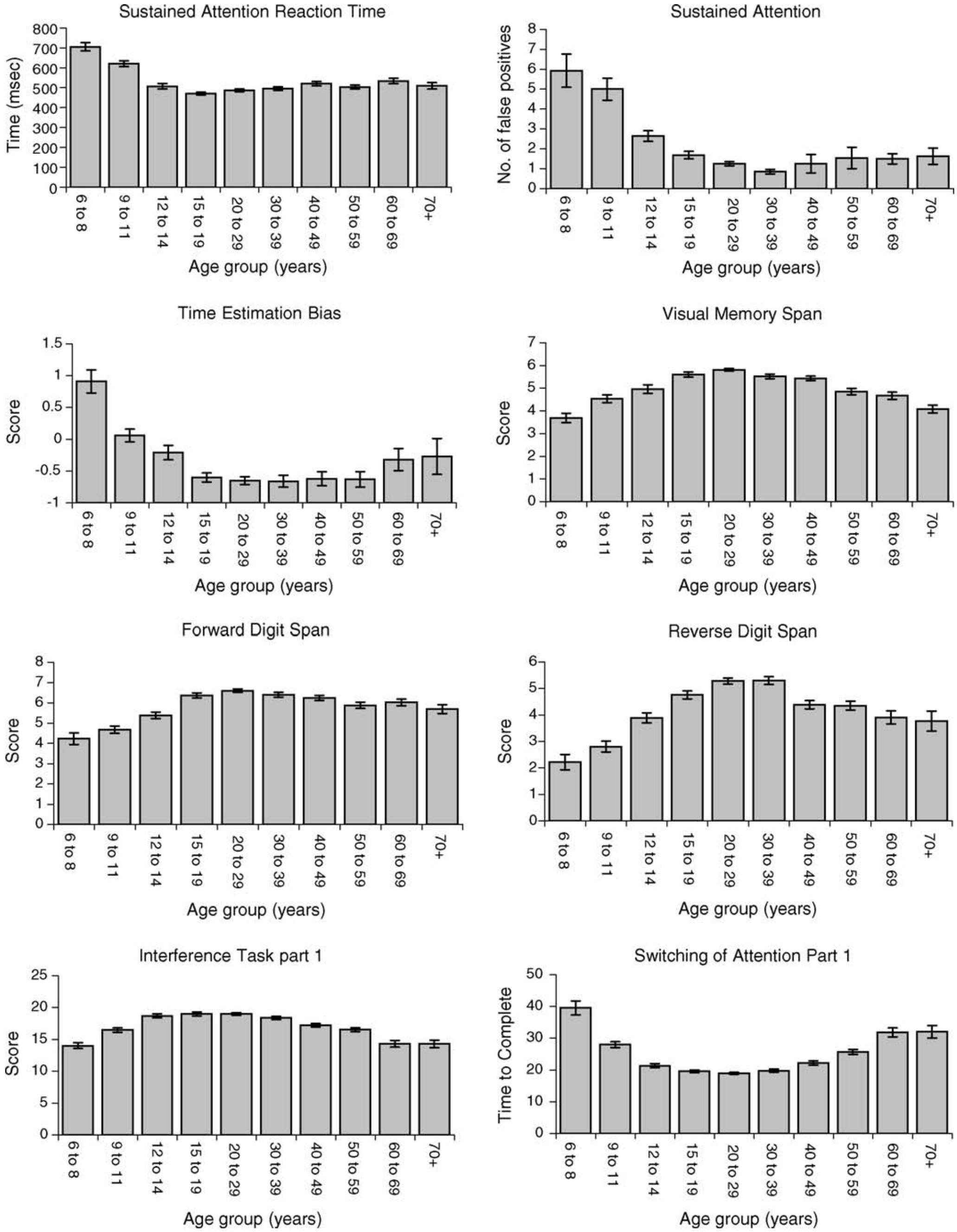

Fig. 1. Means (standard error) by age group on tests of attention and working memory.

which performance gradually decrements. Only Reaction Time asymptotes early (15–19), whilst three measures peak in the 30–39 year age range (i.e. sustained attention target discrimination, Time Estimation and Reverse Digit Span).

2.1.2. Executive function

Performance on four of the six tests of executive function (Verbal Interference Score Part 2, Maze Errors, Maze Overruns, Switching of Attention Score Part 2) was described by a quadratic function over age group with the performance asymptote reached in the 20–29 age group (see Fig. 2). Neither of the measures of self-monitoring from

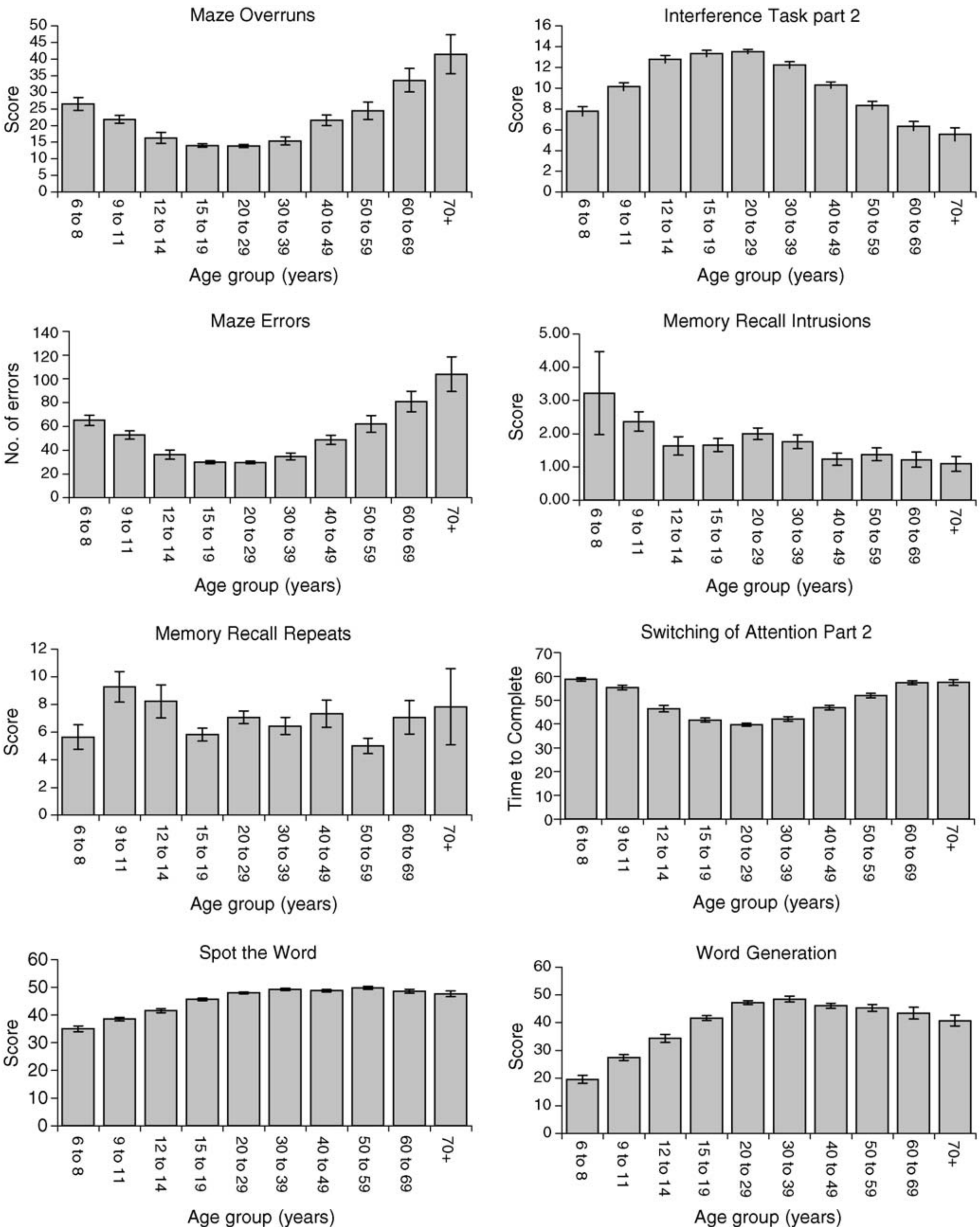

Fig. 2. Means (standard error) by age group on tests of executive function and language.

the Memory Recall test (Intrusions, Repeats) showed robust age effects, despite a small main effect ($p < 0.05$) for Intrusions. The lack of strong effects on these two measures may be related to floor effects over a small scoring range.

2.1.3. Language

A strong quadratic effect of age group was evident for Word Generation (see Fig. 2), with performance increasing linearly through to the 30–39 age group, and declining thereafter. A quadratic effect was also obtained for the Spot the Word Score, though examination of the means over age (Fig. 2) indicates that whilst peak performance is achieved by the 30–39 age group, there is only very minor fall off in performance thereafter.

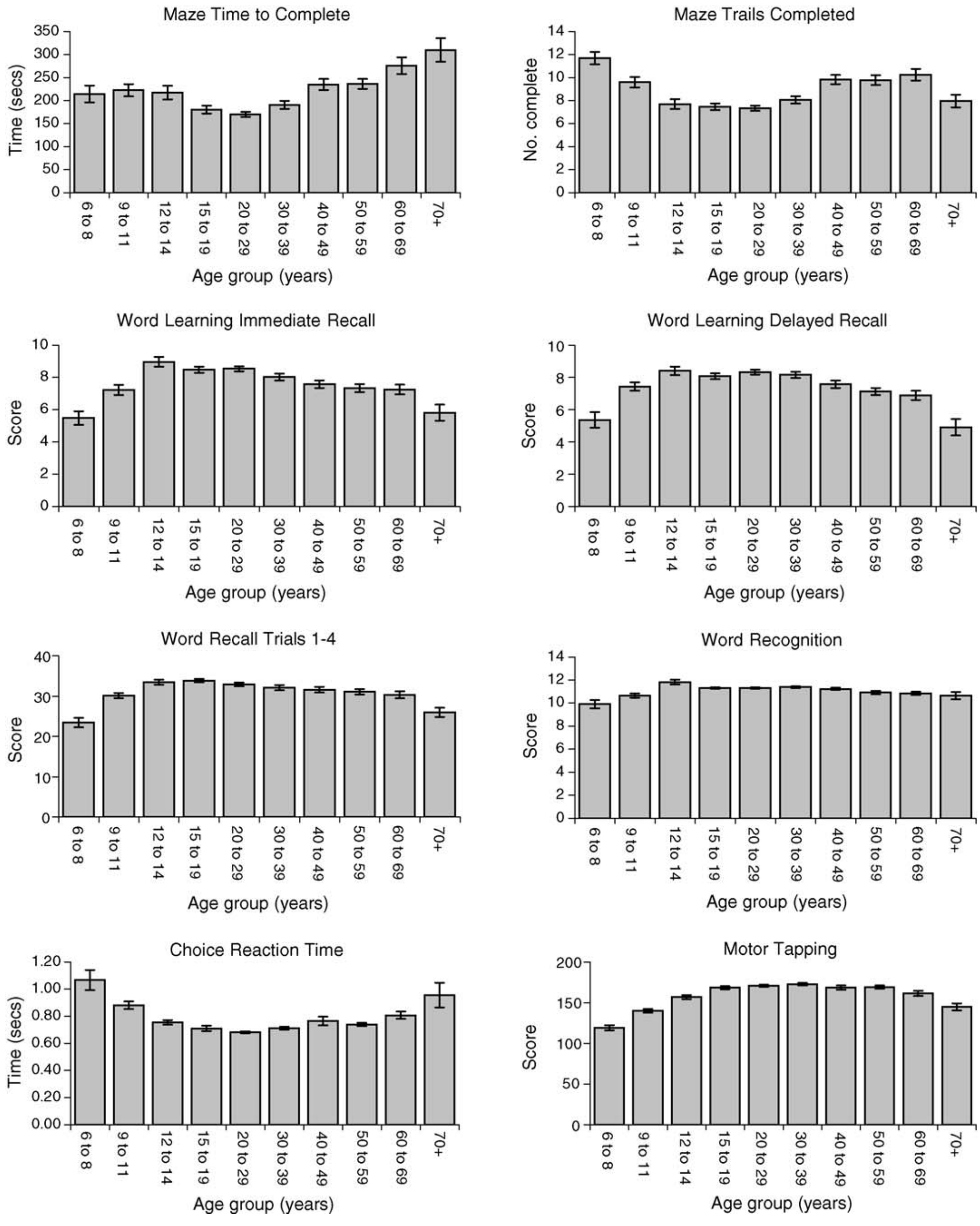

Fig. 3. Means (standard error) by age group on tests of memory and sensori-motor function.

2.1.4. Learning and memory

All six measures were described by a quadratic trend over age group. Measures of verbal learning and memory tended to peak between 12 and 19 years of age, with gradual fall of thereafter (see Fig. 3). By the age of >70, performance level on all four verbal measures approximated that for 6–11 year olds. Learning and memory performance on the maze was also affected quadratically by age group though the trend pattern was more complex. For time to complete (the time taken to learn the maze or until time out), there were only mild improvements between the age range 6 and 29, though with a clear peak in performance in the 20–29 range. Thereafter, performance declined more steeply over the subsequent age ranges. For Maze Trials Completed (the number of trials completed to criterion or until time out), performance improved noticeably through to the 20–29 age group, and declined thereafter through to the 60–69 group.

Table 2
Means and standard deviations by gender on tests of attention and working memory, executive function, language, learning and memory and sensori-motor function

	Gender			
	Males		Females	
	Mean	(S.D.)	Mean	(S.D.)
Attention and working memory				
Time estimation—Bias Score	0.35	1.06	0.51	1.14
Switching of Attention Part 1—Completion Time (s)	23.12	8.27	22.68	8.83
Sustained Attention—False Positives***	2.06	2.77	1.68	3.79
Sustained Attention—Reaction Time (ms)***	501.18	113.21	524.41	112.13
Span of Visual Memory—Score	5.26	1.43	5.21	1.21
Reverse Digit Span—Score	4.49	2.06	4.38	1.82
Interference task—Score Part 1	17.45	3.55	17.93	3.41
Forward Digit Span—Score	6.06	1.60	5.92	1.56
Executive function				
Switching of Attention—Completion Time (s)	46.61	11.60	46.06	11.46
Maze—Error Score***	39.45	41.28	47.01	41.27
Maze—Overruns***	17.22	16.25	20.32	16.00
Interference task—Score Part 2	11.45	4.12	11.28	4.11
Memory Recall and Recognition—Repeats Trials 1–4*	6.93	6.88	7.20	9.14
Memory Recall and Recognition—Intrusions Trials 1–4	2.13	3.77	1.59	2.36
Language				
Word Generation—Words Generated	40.50	13.99	41.74	13.42
Spot the Word—Score	45.67	7.16	46.16	6.56
Learning and memory				
Memory Recall and Recognition—Total Score Trials 1–4**	30.76	6.98	31.78	7.17
Memory Recall and Recognition—Recognition Score	10.96	1.40	11.06	1.30
Memory Recall and Recognition—Immediate Recall Score**	7.52	2.62	8.04	2.60
Memory Recall and Recognition—Delayed Recall Score***	7.34	2.40	7.89	2.52
Maze—Completion Time (s)	201.92	113.45	208.71	109.02
Maze—Trials Completed	7.89	3.56	8.92	4.07
Sensori-motor				
Motor Tapping—Number of Taps (dominant)***	167.00	25.03	158.00	24.49
Choice Reaction Time—Mean Reaction Time (ms)	764.64	250.76	750.90	239.37

* $p \leq 0.05$.
** $p \leq 0.01$.
*** $p \leq 0.001$.

Nominally improved performance in the >70 age group was found to be an artefact of time out effects: an inspection of the scores for this group identified ceiling effects in several subjects who were exceptionally slow at the task and only completed a very small number of trials before the time limit for completion was exceeded. This suggests that the number of trials completed on the task was an inappropriate reference for memory performance in the >70 age group due to the task's time limit.

2.1.5. Sensori-motor

Both measures (Choice Reaction Time, Motor Tapping Score) were described by quadratic trends (see Fig. 3), with performance improving through to the 20–39 age range, and falling thereafter.

2.2. Gender

A main effect for gender was obtained on 9 of the 24 measures (see Tables 1 and 2). On the Sustained Attention task, females made fewer errors but were slower than males. Females made more errors and committed more overruns than males on the Maze. Their recall of verbal information was better than for males, though they made more repeat errors during recall. They made fewer taps than males on the Motor Tapping task.

2.2.1. *Age group by gender*

The main effects obtained for age group and gender were modified by interactions on only three measures: Error rate and Number of Overruns on the Maze and Reaction Time on the Sustained Attention task. In each case, the interaction was explained by an earlier peaking over the age range for females than males on the dependent measure. On the Maze, peak performance for females was obtained in the 15–19 age range, whereas for males the peak was obtained in the 20–29 year range. For Reaction Time on the Sustained Attention task, males peaked in the 15–19 years age group; whilst females peaked in the 12–14 year age range albeit at a marginally slower rate that for males (Fig. 4).

2.2.2. *Years of education*

Level of education affected 7 of the 24 measures assessed. These were Sustained Attention Reaction Time, Forward and Reverse Digit Span, Switching of Attention Part 2 Completion Time, Repeats Score on the Verbal Memory Recall test, Word Generation, Spot the Real Word Score and Time to Complete the Maze (see Tables 1 and 3). All but one of these (Sustained Attention Reaction Time) showed linear improvements in score with increasing years of education. Post hoc analysis of the effects indicated that reaction time during Sustained Attention found no reliable difference between groups.

3. Discussion

The battery used in this study provided sensitive indicators of change in cognition during the years of development of the brain and cognitive system as well as during the years of ageing that follow full development. This outcome was evident for measures of attention and working memory, Word Generation, Verbal Memory Recall, visuospatial learning and memory, executive function and sensori-motor function. In most of these cases, there was improvement in test performance through to some point within the second to fourth decade of life, reflecting the effects of development, followed by gradual decrement or relatively stable performance levels during the subsequent aging years. Age has already been well demonstrated to affect the wide range of higher functions addressed in this study (e.g. Compton et al., 2003; Costa & McCrae, 1993; Rabbitt, 2002; Salthouse, 2000, 2001). The present study, however, considers the full age span in a single study using a single test battery, thereby addressing change over both the developing and declining years.

3.1. *Age*

One observation of interest is the pattern of change in cognitive performance in the younger age groups. In most tests, performance levels incremented in a graded way from the 6 to 8 age band through, in many cases, to the third and fourth decades of life, pointing to a gradual and graded acquisition of skill during the developmental years. There were only few tests in which the rate of cognitive development was markedly greater for very young children, as has been found in some studies (e.g. Anderson et al., 2001; Archibald & Kerns, 1999), and which might be considered attributable to the delayed myelination of tertiary cortical regions. In the present study, however, an asymmetrically greater development in the 6–8 age band relative to the 9–12 age band was found only for measures of Delayed Verbal Recall, simple visuomotor attention (SOA Part 1) and time estimation. Thus, the view that the rate of cognitive development in children under the age of 9 is markedly greater than that in adolescents (Korkman et al., 2002) does not seem to hold in terms of the age bands tested in this study. One test that conformed to the expected effects of delayed myelination was the delay of improvement in Maze Learning Time (Time to Complete) until the early teen years (12–14). The results of this study also provide clear indicators of cognitive ageing, with clear and gradual decline in performance over the decades of late middle to old age. In some cases, however, the decline was only minimal, such as for recognition memory, word knowledge, vigilance speed (Reaction Time during Sustained Attention), memory span (Digit Span), time estimation, verbal fluency and visual discrimination error (Sustained Attention False Positives). In other cases, decline was particularly marked, such as in some measures of executive function (Maze Overruns, Maze Errors, Interference task Part 2) and in some measures of learning and memory (Maze Trials Completed, Delayed Verbal Recall). The study suggests that tests need to be normed over age to help ensure cross-sectional validity in

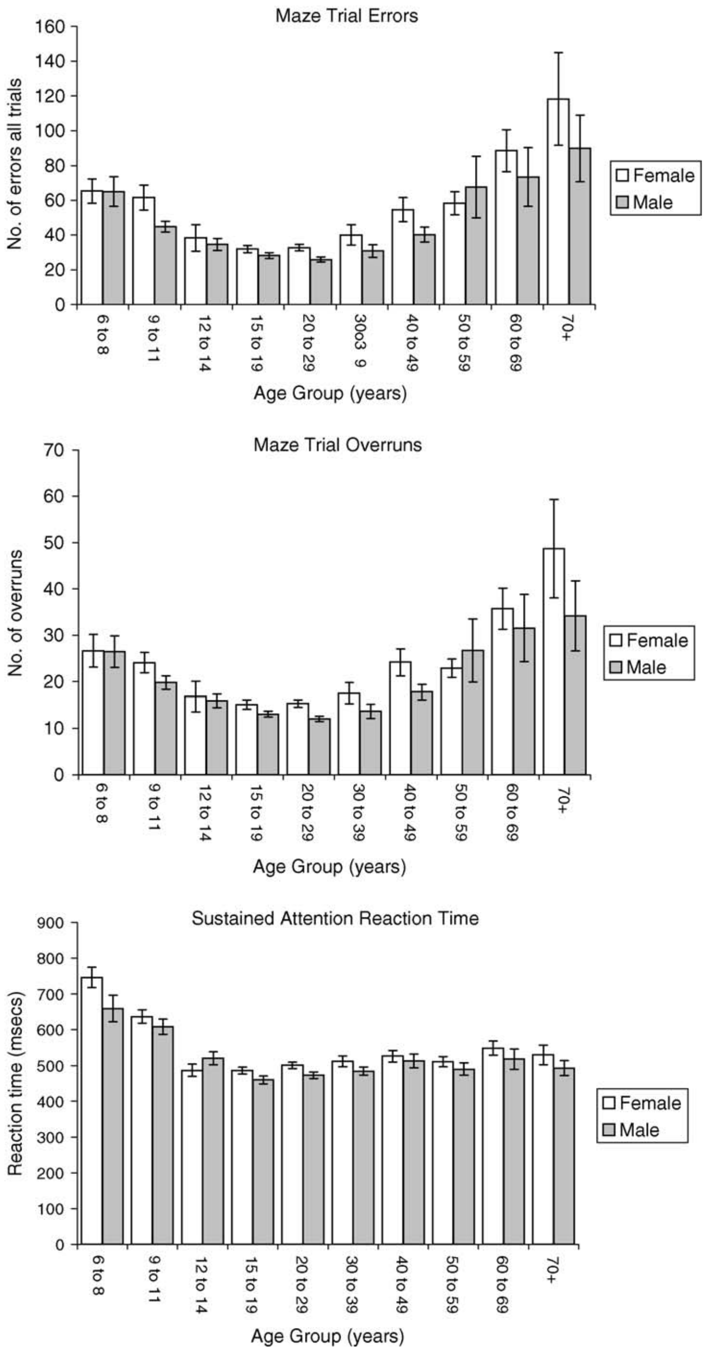

Fig. 4. Means (standard error) by age group and gender on tests of attention and executive function.

clinical assessments, once the utility of this approach has been assessed from appropriate diagnostic and treatment studies. The natural decline in performance with age across many measured functions, as demonstrated in this study, also shows the importance of adequate age norms both during initial assessment in cases of suspected mild cognitive impairment and when following such cases longitudinally over time (see Sarazin & Dubois, 2002). The results from the current study suggest that decline in many cognitive functions from young adulthood into advanced age is quite normal. However, it is likely that some older individuals experience even greater change in cognitive function compared to their peers, such as in cases of mild cognitive impairment and dementia. Our results strongly support the need to

Table 3
Means and standard deviations by years of education (adult participants only) on tests of attention, working memory, executive function, language, learning and memory and sensori-motor function

	Years of education							
	<10		10–12		13–15		>15	
	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)	Mean	(S.D.)
Attention and working memory								
Time estimation—Bias Score	0.07	1.06	−0.51	1.06	−0.65	1.11	−0.62	1.02
Switching of Attention Part 1—Completion Time (s)	27.54	10.73	22.57	8.26	20.65	6.23	21.11	6.66
Sustained Attention—False Positives	3.67	3.96	1.67	3.68	1.29	1.97	1.23	3.04
Sustained Attention—Reaction Time (ms)**	567.67	135.02	487.52	93.95	499.52	100.64	489.87	94.51
Span of Visual Memory—Score	4.59	1.50	5.23	1.32	5.46	1.21	5.59	1.04
Reverse Digit Span—Score***	3.23	1.84	4.29	1.81	4.91	1.64	5.20	1.86
Interference task—Score Part 1	16.63	3.46	17.49	3.76	18.23	3.44	18.30	3.11
Forward Digit Span—Score***	4.98	1.56	5.96	1.38	6.38	1.43	6.58	1.42
Executive function								
Switching of Attention Part 2 Completion Time (s)*	52.51	10.44	46.94	10.52	43.30	11.10	42.98	11.24
Maze—Error Score	49.31	36.27	49.79	55.37	37.91	31.47	39.18	39.02
Maze—Overruns	21.03	14.05	21.56	22.83	16.60	11.20	17.14	14.74
Interference task—Score Part 2	10.35	3.78	11.04	3.99	12.12	4.27	11.86	4.18
Memory Recall and Recognition—Repeats Trials 1–4*	6.36	7.13	7.75	10.84	6.75	9.34	4.76	7.87
Memory Recall and Recognition—Intrusions Trials 1–4	1.43	1.08	1.60	2.20	1.86	2.82	1.49	2.05
Language								
Word Generation—Score**	29.09	12.40	41.51	10.89	45.63	11.38	48.02	11.20
Spot the Word—Score***	38.69	7.32	46.23	5.74	47.94	5.45	50.10	5.39
Learning and memory								
Memory Recall and Recognition—Total Score Trials 1–4	28.27	8.69	30.08	8.18	25.64	10.04	26.51	10.73
Memory Recall and Recognition—Recognition Score	10.97	1.34	10.99	1.46	11.30	1.19	11.32	0.89
Memory Recall and Recognition—Immediate Recall Score	7.29	2.72	7.33	2.30	8.23	2.47	8.03	2.51
Memory Recall and Recognition—Delayed Recall Score	7.00	2.66	7.30	2.17	7.99	2.51	7.95	2.43
Maze—Completion Time*	198.61	94.83	240.60	123.37	194.02	104.38	198.95	112.82
Maze—Trials Completed	9.15	3.97	8.52	3.81	8.02	3.48	8.17	4.08
Sensori-motor								
Motor Tapping—Number of Taps (dominant hand)	144.03	25.34	166.34	21.99	168.54	23.06	169.98	20.66
Choice Reaction Time	874.50	349.73	729.80	159.25	732.11	248.66	699.26	117.20

* $p \leq 0.05$.
** $p \leq 0.01$.
*** $p \leq 0.001$.

represent aging associated cognitive change across a continuum in order to detect sensitively abnormal variation that may occur in such conditions.

One possible concern in the present study is a somewhat skewed distribution of subjects across the age groups. Very young and very old subjects are less easy to acquire than subjects in the middle age bands. As a result, the database tends to contain a relatively larger number of subjects in their twenties and thirties. Fortunately, this age group has reached the full potential in brain functioning and has not yet succumbed to age-related changes that could affect scores on standardized testing. Nevertheless, database numbers are still being increased with an extended effort to elevate numbers and related statistical power in the less well-represented bands. Missing data also presented a minor difficulty, particularly in the youngest age groups but this is not an issue specific to IntegNeuro but rather to the difficulties in testing younger individuals more generally. While this is not so much of an issue within age groups, as variables with large amounts of missing data can be simply excluded from analysis, it can somewhat skew the overall results between age groups.

Another possible concern is that some of the tasks in the battery might have been challenging for the younger children and in particular Verbal fluency, Spot the Real Word, Switching of Attention and Time Estimation. Examination of the scores on these tests for the youngest age group (6 years old) revealed completion rates of 64% on the Switching of Attention task, 95% for Verbal Fluency and 100% for each of Spot the Real Word and Time Estimation. For 6 year olds that did complete these tests, mean scores were one standard deviation above zero for time estimation and over two standard deviations for the other tests. Thus, it appears that the majority of 6 year olds can complete these tests and evidence of only minimal floor effects for this age. But it should be noted that with one exception, the test battery should pick up any child who finds a test too difficult—by virtue of a practice trial protocol that ensures the test is skipped after three practice test failures. The exception is the Switching of Attention task since, unlike the practice trials, which include only short number letter sequences, the test trials presume knowledge of numbers up to 25 and knowledge of the first 13 letters of the alphabet.

A final concern here is the choice of age bands used in analysis. For example, there are substantial changes in cognitive development between 6–8 and 9–11, with normative datasets for children in these ages usually divided into 6-month periods to capture the rapid maturation for children in these age ranges. Clearly, the number of participant datasets available for the study did not permit finer age divisions. However, the aim of the study was to demonstrate the sensitivity of the battery across the age range rather than provide a definitive set of norms for assessment purposes. With respect to the latter issue, we have recently developed a normalization model based on age regression (Crawford & Garthwaite, 2002; Gordon, Cooper, Rennie, Hermens, & Williams, 2005) that provides precise aged-based normative reference for assessment purposes.

3.2. Gender

The present study also examined the effect of gender on the test battery. Females demonstrated better verbal memory than males, though self-monitoring of retrieval context was poorer than for males. Dextral motor expression was generally slower for females. This was also evident during visuospatial learning and memory, where they were also more error prone than males. They were also slower to react during Sustained Attention on verbal material but were more accurate than males. In some tests, there were indications of females reaching an optimal performance level earlier than males. These results are consistent with earlier studies of verbal memory (e.g. Beatty, Mold, & Gontkovsky, 2003; Ruff, Light, & Quayhagen, 1989) and motor dexterity (Brandon, Chavez, & Bennet, 1986; Coleman, Moberg, Ragland, & Gur, 1997), but failed to find evidence for female advantage on executive functioning (Boone, 1999; Boone, Ghaffarian, Lesser, & Hill-Gutierrez, 1993), verbal fluency (Loonstra, Tarlow, & Sellers, 2001; Ruff, Light, & Parker, 1996; Zappala, Measso, Cavarzeran, & Grigoletto, 1995) or visuospatial memory (Caffarra, Vezzadini, Dieci, Zonato, & Venneri, 2002).

Overall, gender effects were obtained on only a third of the measures reported in the present study. The literature contains many studies reporting a relative lack of gender effects on measures of cognitive functioning. This has been demonstrated in the 4–18 age group, where gender has not been found unrelated to verbal fluency (Regard, Strauss, & Knapp, 1982), executive functioning (Paniak, Miller, Murphy, & Patterson, 1996; Rosselli & Ardila, 1993), sensory and motor functioning (Arceneaux et al., 1997), auditory verbal learning and memory (Van den Burg & Kingma, 1999) or visuospatial construction (Fastenau, Denburg, & Hufford, 1999; Fernando, Chard, Butcher, & McKay, 2003). It has also been demonstrated in the adult age range for attention (Beatty et al., 2003), visuospatial construction and memory (Boone, Lesser, Hill-Gutierrez, & Berman, 1993; Coman et al., 2002 Fastenau et al., 1999; Ruff et al., 1996), visuospatial learning (e.g. Ruff et al., 1996), visuomotor tracking and general attention (Giovagnoli, 1997), visual retention and general memory (Zappala et al., 1995), verbal fluency (Loonstra et al., 2001), verbal naming (Cruice, Worrall, & Hickson, 2000; Kent & Luszcz, 2002), word knowledge (Crowell, Vanderploeg, Small, Graves, & Mortimer, 2002), some tests of attention and working memory (Vlahou & Kosmidis, 2002) and on the Halstead-Reitan Neuropsychological test battery (Elias, Robbins, Walter, & Schultz, 1993).

Overall, the gender findings were of modest magnitude. However, it is important to note that the gender literature would suggest that sex differences are expected only on certain types of task that assess gender-sensitive cognitive functions. Sex differences would not be expected a priori on most tests in this battery.

3.3. Years of education

Finally, the present study examined the effect of years of education on cognitive performance. Most noticeable in this regard was the effect of education on measures of verbal fluency (Word Generation), word knowledge (Spot the

Real Word), verbal memory capacity (Forward Digit Span), verbal working memory (Reverse Digit Span), mental adaptability (SOA Time Part 2), verbal self-monitoring (Verbal Recall Repeats) and visuospatial learning and memory (Maze Completion Time). The effects of years of education on a range of measures of language is perhaps not surprising given the focus of educational systems on linguistic abilities (e.g. Ardila, Ostrosky-Solis, Rosselli, & Gomez, 2000).

An effect of education on cognitive function has often been reported (e.g. Ardila et al., 2000; Bravo & Hebert, 1997; Collie, Shafiz-Antonacci, Maruff, Tyler, & Currie, 1999; Gladsjo et al., 1999; Loonstra et al., 2001; Richardson & Marottoli, 1996; Vlahou & Kosmidis, 2002) though it must be considered that any such relationship would be susceptible to confound by other factors, such as age, intelligence and gender. Ardila et al. (2000), for example, found education effects to be noticeably greater than those of age on a neurocognitive battery assessment of orientation, attention, memory, language, visuoperceptual abilities, motor skills and executive functions. At the same time, there have been studies that reported no effects or only minimal effects of years of education (e.g. Boone, 1999; Fastenau et al., 1999; Jones & Gallo, 2001; Meguro et al., 2001).

The present study differs significantly from many earlier studies of the effect of years of education on cognitive function and this may go some way to explaining some of the differences observed. Firstly, a number of investigations have restricted the age range examined to those 60 years or older, with the specific aim of identifying factors relevant to the onset of age-related cognitive dysfunction (e.g. Bravo & Hebert, 1997; Gontkovsky, Mold, & Beatty, 2002; Rapp, Espeland, Hogan, Jones, & Dugan, 2003; Richardson & Marottoli, 1996). This contrasts markedly with the present study, which examined the full age range. Secondly, several studies have employed only a limited range of measures (e.g. Vlahou & Kosmidis, 2002) or limited testing to only one gender (e.g. Rapp et al., 2003). Again, the present study examined a wide range of measures. Thirdly, assessments of education effects vary in the categorization of level of education. For example, a number of reports have been based on participants with relatively low levels of education (e.g. Ardila et al., 2000; Laks et al., 2003). In the study by Ardila et al. (2000), three of their four education level categories employed involved participants with less than 10 years of education. In contrast, three of the four levels of education categories in the present study encompassed those with 10 years of education or more, and few of those with less than 10 years of education had been at school for less than 8 years. A number of other studies where participant selection has been biased more towards the upper end of the education scale (e.g. Richardson & Marottoli, 1996), which tends to be more strongly populated in western acculturated regions, education has been found to explain only minimal data variance.

3.4. Standardization

The delivery of the best battery in the present study involved standardized conditions in which human contact during testing was minimized, in which test instruction is delivered using pre-recorded audio and video in a two-way interactive setting and in which sound, light and other environmental factors are prescribed. It is suggested that such requirements go a long way towards minimizing variability in test performance that can result across settings due to factors, such as tester bias, variation in test instruction delivery, pace of test delivery (e.g. duration of stimulus presentation and inter-stimulus delays). Nevertheless, some of these same requirements may be seen as a source of weakness when applied to clinical populations. One comment raised often in regard to computerized test batteries is that neuropsychological screening should be supervized by a testing psychologist. The argument is that the psychologist should be able to observe client behavior during test performance, just in case this information might be qualitatively useful in explaining abnormalities in test performance. This would particularly be the case with older groups that might suffer more physical disabilities and who may have more difficulty pressing a touchscreen to register responses.

These are important issues in the case where the test battery used in the present study is applied to the clinical setting. However, the issues are well accommodated by the standardized approach adopted. For example, video monitoring permitted clear observation of test behavior and audio communication allowed contact when and as required. Also, the mode of delivery of instructions was fail-safe and provided an attentional cueing mechanism for test supervision. During instruction delivery, embedded software monitored the client's behavior to ensure that the relatively low demand test trials were performed correctly from a behavioral (i.e. correct response action) and a cognitive (i.e. correct response content) standpoint. Three consecutive failures resulted in a warning being sent to the test supervisor to take note of the failure and to initiate appropriate action with the client. Information about whether a client fails test instructions on none, one, two or three occasions provides useful information analogous to qualitative observation. Further, the highly sensitive, computerized touchscreen allowed quantitative measurement of response behavior. For example, the touchscreen system used allowed measurement of response variability, such as the standard deviation of the inter-tap interval on the Motor Tapping task. This informs quantitatively about difficulties with motor control, analogous to the qualitative information the neuropsychologist might obtain by observing tapping behavior, but better since the availability of the normative database described in this paper permits statistical assessment of the abnormality of such measures.

3.5. Summary

This study provides insight into the changes of cognitive functions though from the earliest years to the beginnings of decline, and identifies the concomitant contributions of gender and years of education. The test battery that is used is shown to provide a sensitive assessment of cognitive function across the full age range. In this regard, the test battery and its associated database has broad applicability in areas, such as neuropsychological assessment, assessment of treatment effects, longitudinal case management and the assessment of drug efficacy. An important future development using the battery is the collection of data from a wide range of psychopathologies, to permit the development of wide-ranging specificity analysis.

Acknowledgement

We acknowledge and thank the Brain Resource Company (www.brainresource.com) for permission to use the Brain Resource International Database.

References

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

Anderson, P., Anderson, V. A., Northam, E., Jacobs, R., & Catroppa, C. (2001). Development of executive functions through late childhood and adolescence in an Australian sample. *Developmental Neuropsychology*, *20*, 385–406.

Arceneaux, J. M., Hill, S. K., Chamberlin, C. M., & Dean, R. S. (1997). Developmental and sex differences in sensory and motor functioning. *International Journal of Neuroscience*, *89*, 253–263.

Archibald, S. J., & Kerns, K. A. (1999). Identification and description of new tests of executive functioning in children. *Child Neuropsychology*, *5*, 115–129.

Ardila, A., Ostrosky-Solis, F., Rosselli, M., & Gomez, C. (2000). Age related cognitive decline during normal aging: The complex effect of education. *Archives Clinical Neuropsychology*, *15*, 495–513.

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.

Beatty, W. W., Mold, J. W., & Gontkovsky, S. T. (2003). RBANS performance: Influences of gender and education. *Journal of Clinical & Experimental Neuropsychology*, *25*, 1065–1069.

Benton, A. L., & Hamsher, K. (1989). *Multilingual aphasia examination*. Iowa City, Iowa: AJA Associates.

Boone, K. B. (1999). Neuropsychological assessment of executive functions: Impact of age, education, gender, intellectual level, and vascular status on executive test scores. In B. L. Miller & J. L. Cummings (Eds.), *The human frontal lobes: Functions and disorders. The science and practice of neuropsychology series* (pp. 247–260). New York: Guilford Press.

Boone, K. B., Ghaffarian, S., Lesser, I. M., Hill-Gutierrez, E., et al. (1993). Wisconsin Card Sorting test performance in healthy, older adults: Relationship to age, sex, education, and IQ. *Journal of Clinical Psychology*, *49*, 54–60.

Boone, K. B., Lesser, I. M., Hill-Gutierrez, E., Berman, N. G., et al. (1993). Rey-Osterrieth complex figure performance in healthy, older adults: Relationship to age, education, sex, and IQ. *Clinical Neuropsychologist*, *7*, 22–28.

Brabeck, M. M. (1984). Longitudinal studies of intellectual development during adulthood: Theoretical and research models. *Journal of Research & Development in Education*, *17*(3), 12–27.

Brandon, A. D., Chavez, E. L., & Bennet, T. L. (1986). A comparative evaluation of two neuropsychological finger tapping instruments: Halstead-Reitan and Western Psychological Services. *Clinical Neuropsychology*, *18*, 64–65.

Bravo, G., & Hebert, R. (1997). Age- and education-specific reference values for the Mini-Mental and Modified Mini-Mental State Examinations derived from a non-demented elderly population. *International Journal of Geriatric Psychiatry*, *12*, 1008–1018.

Caffarra, P., Vezzadini, G., Dieci, F., Zonato, F., & Venneri, A. (2002). Rey-Osterrieth complex figure: Normative values in an Italian population sample. *Neurological Sciences*, *22*, 443–447.

Coleman, A., Moberg, P., Ragland, J., & Gur, R. (1997). Comparison of the Halstead-Reitan and Infrared Light Beam Finger Tappers. *Assessment*, *4*, 277–286.

Collie, A., Shafiz-Antonacci, R., Maruff, P., Tyler, P., & Currie, J. (1999). Norms and the effects of demographic variables on a neuropsychological battery for use in healthy ageing Australian populations. *Australian & New Zealand Journal of Psychiatry*, *33*, 568–575.

Coman, E., Moses, J. A. J. R., Kraemer, H. C., Friedman, L., Benton, A. L., & Yesavage, J. (2002). Interactive influences on BVRT performance level: Geriatric considerations. *Archives of Clinical Neuropsychology*, *17*, 595–610.

Compton, D., Avet-Compton, T., Bachman, L., & Brand, D. (2003). Working memory and perceptual speed mediating of age-associated changes in cognition within a sample of highly-educated adults. *North American Journal of Psychology*, *5*, 451–477.

Costa, P. T., & McCrae, R. R. (1993). Psychological research in the Baltimore longitudinal study of aging. *Zeitschrift fur Gerontologie*, *26*, 138–141.

Crawford, J. R., & Garthwaite, P. H. (2002). Investigation of the single case in neuropsychology: Confidence limits on the abnormality of test scores and test score differences. *Neuropsychologia*, *40*, 1196–1208.

Crowell, T. A., Vanderploeg, R. D., Small, B. J., Graves, A. B., & Mortimer, J. A. (2002). Elderly norms for the Spot-the-Word test. *Archives of Clinical Neuropsychology*, *17*, 123–130.

Cruice, M. N., Worrall, L. E., & Hickson, L. M. H. (2000). Boston Naming test results for healthy older Australians: A longitudinal and cross-sectional study. *Aphasiology*, *14*, 143–155.

Elias, M. F., Robbins, M. A., Walter, L. J., & Schultz, N. R. (1993). The influence of gender and age on Halstead-Reitan Neuropsychological test performance. *Journal of Gerontology*, *48*, 278–281.

Fastenau, P. S., Denburg, N. L., & Hufford, B. J. (1999). Adult norms for the Rey-Osterrieth Complex Figure test and for supplemental recognition and matching trials from the Extended Complex Figure test. *Clinical Neuropsychologist*, *13*, 30–47.

Fernando, K., Chard, L., Butcher, M., & McKay, C. (2003). Standardization of the Rey Complex Figure test in New Zealand children and adolescents. *New Zealand Journal of Psychology*, *32*, 33–38.

Geffen, G. M., Moar, K. J., Hanlon, A. P., Clark, C. R., & Geffen, L. B. (1990). Performance measures of 16- to 86-year-old males and females on the Auditory Verbal Learning test. *The Clinical Neuropsychologist*, *4*, 45–63.

Gilhooly, K. L., & Logie, R. H. (1980). Age of acquisition, imagery, concreteness, familiarity and ambiguity measures for 1944 words. *Behaviour Research Methods and Instrumentation*, *12*, 395–427.

Giovagnoli, A. R. (1997). Trail Making test: Normative values from 287 normal adult controls. *Italian Journal of Neurological Sciences*, *17*, 305–310.

Gladsjo, J. A., Schuman, C. C., Evans, J. D., Peavy, G. M., Miller, S. W., & Heaton, R. K. (1999). Norms for letter and category fluency: Demographic corrections for age, education, and ethnicity. *Assessment*, *6*, 147–178.

Gontkovsky, S. T., Mold, J. W., & Beatty, W. W. (2002). Age and educational influences on RBANS index scores in a nondemented geriatric sample. *Clinical Neuropsychologist*, *16*, 258–263.

Gordon, E. (2003a). Integrative neuroscience and psychiatry. *Neuropsychopharmacology*, *28*(Suppl. 1), S2–S8.

Gordon, E. (2003b). Integrative neuroscience in psychiatry: The role of a standardized database. *Australasian Psychiatry*, *11*(2), 156–163.

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

Halstead, W. C. (1947). *Brain and intelligence*. Chicago: University of Chicago Press.

Hickie, I. (1998). SPHERE: A National Depression Project. *Australasian Psychiatry*, *6*, 248–250.

Jones, R. N., & Gallo, J. J. (2001). Education bias in the mini-mental state examination. *International Psychogeriatrics*, *13*, 299–310.

Kent, P. S., & Luszcz, M. A. (2002). A review of the Boston Naming test and multiple-occasion normative data for older adults on 15-item versions. *Clinical Neuropsychologist*, *16*, 555–574.

Korkman, M., Kemp, S. L., & Kirk, U. (2002). Effects of age on neurocognitive measures of children ages 5 to 12: A cross-sectional study on 800 children from the United States. *Developmental Neuropsychology*, *20*, 331–354.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence: Brown University Press.

Laks, J., Batista, E. M. R., Guilherme, E. R. L., Contino, A. L. B., Faria, M. E. V., Figueira, I., et al. (2003). Mini-mental state examination in community-dwelling elderly: Preliminary data from Santo Antonio de Padua, Rio de Janeiro, Brazil. *Arquivos de Neuro-Psiquiatria*, *61*, 782–785.

Lezak, M. (2004). *Neuropsychological assessment*. New York: OUP.

Loonstra, A. S., Tarlow, A. R., & Sellers, A. H. (2001). COWAT metanorms across age, education, and gender. *Applied Neuropsychology*, *8*, 161–166.

Meguro, K., Ishii, H., Yamaguchi, S., Ishizaki, J., Sato, M., Hashimoto, R., et al. (2004). Prevalence and cognitive performance of clinical dementia rating 0.5 and mild cognitive impairment in Japan: The Tajiri Project. *Alzheimer Disease & Associated Disorders*, *18*, 3–10.

Meguro, K., Shimada, M., Yamaguchi, S., Ishizaki, J., Ishii, H., Shimada, Y., et al. (2001). Cognitive function and frontal lobe atrophy in normal elderly adults: Implications for dementia not as aging-related disorders and the reserve hypothesis. *Psychiatry & Clinical Neurosciences*, *55*, 565–572.

Milner, B. (1970). Memory and the medial temporal regions of the brain. In K. H. Pribram & D. E. Broadbent (Eds.), *Biology of memory*. New York: Academic Press.

Morrison, P. R., & Gates, G. R. (1988). Assessment of a microcomputer-based version of the Austin Maze. *Journal of General Psychology*, *115*, 307–314.

Paniak, C., Miller, H. B., Murphy, D., & Patterson, L. (1996). Canadian developmental norms for 9 to 14 year-olds on the Wisconsin Card Sorting test. *Canadian Journal of Rehabilitation*, *9*, 233–237.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery and meaningfulness values for 925 words. *Journal of Experimental Psychology Monograph*, *76*(Suppl. (3, Pt. 2)).

Paul, H. R., Gunstad, J., Cooper, N., Williams, M. L., Clark R., Cohen, A. R., Gordon, E. (2006). Cross cultural assessment of neuropsychological performance and electrical brain function measures: additional validation of an international brain database. *International Journal of Neuroscience*, in press.

Paul, R. H., Lawrence, J., Williams, L. M., Clark, C. R., Cooper, N., & Gordon, E. (2005). Preliminary validity of IntegNeuro: A new computerized battery neurocognitive tests. *International Journal of Neuroscience*, *115*(11), 1549–1567.

Rabbitt, P. (2002). Aging and cognition. In H. Pashler & J. Wixted (Eds.), *Steven's handbook of experimental psychology, vol. 4: Methodology in experimental psychology* (3rd ed., pp. 793–860). New York, NY, US: John Wiley & Sons, Inc, x, p. 901.

Rapp, S. R., Espeland, M. A., Hogan, P., Jones, B. N., & Dugan, E. (2003). Baseline experience with Modified Mini Mental State Exam: The Women's Health Initiative Memory Study (WHIMS). *Aging & Mental Health*, *7*, 217–223.

Regard, M., Strauss, E., & Knapp, P. (1982). Children's production on verbal and non-verbal fluency tasks. *Perceptual & Motor Skills*, *55*, 839–844.

Reitan, R. M. (1955). The relation of the Trail Making test to organic brain damage. *Journal of Consultation Psychology*, *5*, 393–394.

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

Rey, A. (1964). *L'examen clinique en psychologie*. Paris: Presses Universitaires de France.

Richardson, E. D., & Marottoli, R. A. (1996). Education-specific normative data on common neuropsychological indices for individuals older than 75 years. *Clinical Neuropsychologist*, *10*, 375–381.

Rosselli, M., & Ardila, A. (1993). Developmental norms for the Wisconsin Card Sorting test in 5- to 12-year-old children. *Clinical Neuropsychologist*, *7*, 145–154.

Roth, D. L., & Crosson, B. (1985). Memory Span and long-term memory deficits in brain-injured patients. *Journal of Clinical Psychology*, *41*, 521–527.

Ruff, R., Light, R., & Parker, S. (1996). Visuospatial learning: Ruff Light Trail Learning test. *Archives of Clinical Neuropsychology*, *11*, 313–327.

Ruff, R. M., Light, R. H., & Quayhagen, M. (1989). Selective Reminding tests: A normative study of verbal learning in adults. *Journal of Clinical & Experimental Neuropsychology*, *11*, 539–550.

Salthouse, T. (2000). Aging and measures of processing speed. *Biological Psychology*, *54*, 35–54.

Salthouse, T. (2001). Structural models of the relations between age and measures of cognitive functioning. *Intelligence*, *29*, 93–115.

Sarazin, M., & Dubois, B. (2002). Mild cognitive impairment or pre-demential Alzheimer's disease? *Revista de Neurologia*, *158*(10 Suppl.), S30–S34.

Herb, S. H. (1995). Children and learning. In N. L. Herron & D. Zabel (Eds.), *Bridging the gap: Examining polarity in America* (pp. 197–236). CO, US: Englewood.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.

Sullivan, R., Senior, G., & Hennessy, M. (2000). Australian age-education and premorbid cognitive/intellectual estimates for the WAIS-III. In *Poster presented at the 6th Annual Conference of the APS College of Clinical Neuropsychologists, NSW, Australia*.

Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms*. New York: Erlbaum.

Van den Burg, W., & Kingma, A. (1999). Performance of 225 Dutch school children on Rey's Auditory Verbal Learning test (AVLT): Parallel test–retest reliabilities with an interval of 3 months and normative data. *Archives of Clinical Neuropsychology*, *14*, 545–559.

Vlahou, C. H., & Kosmidis, M. H. (2002). The Greek Trail Making test: Preliminary normative data for clinical and research use. *Psychology: The Journal of the Hellenic Psychological Society*, *9*, 336–352.

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

Williams, L. J., Paul, R. H., Lawrence, J., Clark, C. R., Cooper, N., & Gordon, G. (2005). The test–retest reliability a Standardized Neurocognitive and Neurophysiological test battery: "Neuromarker". *International Journal of Neuroscience*, *115*(12), 1605–1630.

Zappala, G., Measso, G., Cavarzeran, F., & Grigoletto, F. (1995). Aging and memory: Corrections for age, sex and education for three widely used memory tests. *Italian Journal of Neurological Sciences*, *16*, 177–184.