Using neuropsychological test performance to identify Alzheimer's disease at a South African memory clinic

Megan Nield

Department of Psychology
University of Cape Town, South Africa

Supervisor: Kevin G. F. Thomas, Ph.D.

Word count: 7738

ABSTRACT

Cross-cultural effects on neuropsychological and cognitive test performance may mean that normative data obtained from North American and European samples do not apply to the South African population. There is a need for normative studies on the South African population in order to effectively test individuals on, for instance, neuropsychological tests such as those used to identify Alzheimer's disease (AD). This study describes the performance of three groups of South African individuals on the battery of neuropsychological tests routinely used at Groote Schuur Hospital's Memory Clinic. Participants were 29 healthy volunteer young adults (aged 18 to 25 years); 25 healthy volunteer older adults (65 years and older); and 24 Memory Clinic patients (65 years and older) with an independent diagnosis of possible/probable AD. All participants were administered the CLOX executive clock drawing task, verbal fluency and category naming tests, the Rey Auditory-Verbal Learning Test, the Rey Osterreith Complex Figure Test, tests of forward and backward digit span, and the Trail Making Test. Results show that this test battery is useful in discriminating between healthy and demented older adults in South Africa. Further, this study provides a first step toward collecting normative data for older South African adults on some of the most commonly-used memory clinic neuropsychological tests.

Keywords: South African; Alzheimer's Disease; memory clinic; elderly; cross-cultural; neuropsychological assessment.

Performance on neuropsychological tests has been shown to be influenced by individual differences on numerous biological and socio-cultural variables, separate from clinical factors. Biological variables include age, gender, handedness, and perhaps IQ. Socio-cultural variables include race/ethnicity, level and quality of education, socio-economic status, level of acculturation, and language variables, including deviance from English as first language and bilingualism (Boone, Victor, Wen, Razani, & Ponton, 2007; Lezak et al., 2004; Mitrushina, Boone, Razani, & D'Elia, 2005; Nell, 2000).

Particularly for neuropsychologists operating in memory clinics, where the typical patient is older than 65 years, it is imperative to consider the effects of age on neuropsychological test performance. Even in the absence of disease processes and related cognitive impairment, advancing sensory and motor difficulties in elderly populations adversely affect such test performance, making it is easy to draw erroneous conclusions about the neuropsychological status of the individual (Lezak et al., 2004). Norms for neuropsychological tests therefore need to be adjusted when applied to older adults to ensure, for instance, more accurate detection of cognitive impairments associated with both normal aging and with age-related diseases such as dementia of the Alzheimer's type (AD; Lezak et al., 2004; Steinberg, Bieliauskas, Smith, Langelloti, et al., 2005).

Alzheimer's disease is a neurodegenerative condition that progressively affects a widening range of cognitive functions, particularly memory. Declining cognitive function is associated with degeneration of nervous tissue resulting from the accumulation of neurofibrillary tangles and amyloid plaques (Kandel, Schwartz, & Jessell, 2000; Lezak, Howieson & Loring, 2004). The main regions affected are the hippocampus, temporal lobes, prefrontal and parietal areas, all of which are critical to memory functioning. Thus, memory impairment usually occurs early in the course of the disease, with impairment of other cognitive functions (e.g., problem solving, attention, language, calculation, and visuospatial abilities) spreading gradually and more generally over time. In order to detect the early onset of the disease, neuropsychologists in memory clinics administer batteries of neuropsychological tests, which measure patient

deviations from standardized norms (Gaither, 2007; Hodges, 2001; Kandel et al., 2000; Lezak et al., 2004).

In the context of South African memory clinics and neuropsychology, the effects of age need to be considered alongside the marked individual differences in IQ, quality of education, and a range of other socio-cultural factors. Steinberg and Bieliauskas (2005) found that IQ accounts for a larger proportion of variance in performance on neuropsychological tests than does typically measured socio-cultural variables (e.g., age or education). They propose that this may be because performances on IQ tests and neuropsychological tests are both normally distributed. Additionally, most IQ tests measure similar cognitive functions as are assessed in neuropsychological tests, and so there may be some shared variance. IQ adjustments can be made on actual test performances; compared with imprecise and problematic education measurements,[1] they are more reliable for fair application of tests to different individuals (Steinberg, Bieliauskas, Smith, Langelloti, et al., 2005).

In the Mayo Clinic Older Americans Normative Studies (MOANS), neuropsychological tests were administered to a reasonably homogeneous sample of adults aged 56 to 99 years. These individuals were English-speaking, Caucasian, of above average intelligence, and from the same geographic region in the United States (Steinberg & Bieliauskas, 2005; Steinberg, Bieliauskas, Smith, & Ivnik, 2005; Steinberg, Bieliauskas, Smith, Ivnik, et al., 2005; Steinberg, Bieliauskas, Smith, Langelloti, et al., 2005). These studies were aimed at developing age- and IQ-adjusted norms for performance on a range of neuropsychological tests.

The MOANS series found that full scale IQ scores (measured using the Wechsler Adult Intelligence Scale - Revised; WAIS-R; Wechsler, 1981) were more strongly related to age-adjusted test performance scores, than was education level. The authors therefore concluded that adjusting for IQ as well as age may be most appropriate for the application of these tests to older populations. They acknowledge that the universal application of their age- and IQ-adjusted norms to ethnically, racially, intelligently, and geographically different populations may not be entirely appropriate, but suggest, due to the enormity of their sample, that the norms provided

still be used as a guideline for future clinical and research purposes (Steinberg, Bieliauskas, Smith, & Ivnik, 2005).

Although these MOANS norms can be useful, in applying them to South African clinical and research studies they may need to be adjusted for race, ethnicity, and other cultural and geographical factors of relevance to this context. The South African population comprises many different race and ethnic groups. Boone et al. (2007), in a North American cross-cultural study, found significant differences in neuropsychological test performance between different ethnic groups. Specifically, they found that Caucasians, compared to other ethnic groups, obtained higher scores on many of the administered tests. The authors concluded that members of these other ethnic groups were therefore more likely to be 'overpathologized', and they therefore questioned the usefulness of Caucasian-derived normative data in testing other ethnic groups. In studies such as this, where clear differences were found in the neuropsychological test performance of different ethnic and racial groups, race/ethnicity is most probably a proxy for other between-group differences such as quality and level of education, socioeconomic status, and culture.

Under Apartheid government, non-Caucasian people, such as Blacks, Coloureds, Indians, and Asians, were deprived of the right to obtain high-quality education. This means that in the present elderly population of South Africa, a majority of individuals have not had the privileged educational background available to whites of the same generation. For this reason is it useful to consider both the number of years of education, as well as the quality of education received, when interpreting neuropsychological test performance (Nell, 2000; Shuttleworth-Edwards et al., 2004). A low level or quality of education has been shown to be associated with poorer performance on neuropsychological tests, and with a higher risk for AD (Shuttleworth-Edwards et al., 2004; Steinberg, Bieliauskas, Smith, Langelloti, et al., 2005). Interestingly, Shuttleworth-Edwards et al. (2004) showed that quality of education had a greater effect than race on neuropsychological test performance: Black South Africans who had been exposed to advantaged educational backgrounds performed better on neuropsychological tests than those of the same race who were from disadvantaged backgrounds. Future normative studies on the South

African population will have to take into account the mixture of educational backgrounds and the influence these have on neuropsychological test performance.

Culture is possibly the most interesting factor affecting neuropsychological test performance. It impacts on neuropsychological test performance because it shapes the way we think, act and feel (Ardila, 1996). Nervous system development is influenced by life experiences, so that individuals from different cultural and language groups may have acquired different knowledge, and may deal with verbal, non-verbal, and visuospatial cognitive tasks differently (Boone et al., 2007; Ostrosky-Solis & Oberg, 2006; Roselli & Ardila, 2003; Shuttleworth-Edwards et al., 2004). Furthermore, the neuropsychological tests we use today were developed by highly-educated Caucasians, and focus on skills that are highly valued within "white, Western, middle class culture" but that may not necessarily be as highly valued in African cultures (Manly, 2005, p. 272). Indeed, neuropsychological test performance of individuals from non-Western cultures, including African-, Hispanic-, and Asian-based cultures, differs from the performance of Western individuals who participated in the original normative studies of commonly used tests (Mitrushina et al., 2005). According to Shuttleworth-Edwards et al. (2004), the application of "western IQ tests to…African-based cultures is considered especially problematic" (p. 905). Because many of the Western tests used in memory clinics for the detection of AD are similar to tests in IQ batteries, it is critical to consider cross-cultural effects when using those tests in a South African memory clinic.

It may also be useful to consider race/ethnicity and socio-cultural background as separate factors, particularly in the South African context, where there is a mixture of white Western, Asian and African-based cultures. Post-apartheid, previously isolated and disadvantaged people groups are now more likely to be more exposed to Western cultural influences, and are therefore more likely to show acculturation to Western values (Boone et al., 2007; Shuttleworth-Jordan, 1995). Consequently, members of the same race can now have very different learning experiences, socioeconomic status, quality of education, proficiency in English, and levels of bilingualism, depending on the extent of the individual's exposure to Western culture (Shuttleworth-Edwards, 2004; Shuttleworth-Jordan, 1995).

Even considering the substantial literature reviewed above, Ostrosky-Solis and Oberg (2006) claim that "little is known about the cross-cultural aspects of even the most commonly used neuropsychological tests" (p. 321). Some recent studies have, however, focused on cross-cultural effects on specific tests used to detect AD. These studies are reviewed below.

Cross Cultural Effects on the Mini-Mental State Examination (MMSE)

Age and socio-cultural differences affect performance on the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975). For example, in one cross-cultural study, Mexican-American elderly performed more poorly than did non-Mexican Americans; and in another, AD patients from the United Kingdom performed more poorly than did those from the United States (Black et al., 1999; Gibbons et al., 2002). There are no published research studies documenting the performance of South Africans diagnosed with AD on the MMSE.

Cross Cultural Effects on the CLOX

Royall et al. (2003) applied a Spanish translation of the CLOX (Royall, Cordes & Polk, 1998) to a large sample (1309 participants) of Mexican-American older adults, stratified by their degree of North American acculturation. Cross-cultural effects were significant for both CLOX1 and CLOX2, even when age, gender, income, education, and the language of presentation (English or Spanish) were controlled. Specifically, Spanish-speaking participants obtained significantly lower CLOX1 scores than did English-speaking participants. The authors concluded that the significant cultural and language effects found were not important, however, and that valid administration of the Spanish translation of CLOX was indeed possible "regardless of acculturation and education" (Royall, 2003, p. 135). Because the authors were also the original developers of the instrument, it was clearly in their interest to assert its cross-cultural validity, despite the conflicting evidence that was presented. There are no published research studies documenting the performance of South Africans diagnosed with AD on the CLOX.

Cross Cultural Effects on Verbal Fluency

Oberg and Ramirez (2006) compared the phonemic fluency performance of healthy adults from five different countries (Mexico, Argentina, Denmark, USA and Israel) who spoke four different languages (Spanish, Danish, English and Hebrew). They found that when the frequency of the

letter in the subject's language and years of education were taken into account, subjects produced relatively equal numbers of words, regardless of cultural or language background. Although these data suggest that phonemic verbal fluency tests may be culture-fair, only two of the six studies included participants aged 56 years or older. There are no published research studies documenting the performance of South Africans diagnosed with AD on verbal fluency.

Cross Cultural Effects on the Rey Auditory Verbal Learning Test (AVLT)

Estevez et al. (2003) administered the AVLT (Schmidt, 1996, as cited in Lezak et al., 2004) to elderly Spanish patients with subjective memory complaints and followed them longitudinally for 2 years. They found that those patients who were diagnosed with probable AD at the 2 year follow-up had, at first administration, worse scores on measures of learning and delayed recall, and a higher percentage of forgetting. This cross-cultural study suggests that, in a memory clinic type setting, the AVLT is useful for predicting probable AD in a non-western culture. There are no published research studies documenting the performance of South Africans diagnosed with AD on the AVLT.

Cross Cultural Effects on the Rey-Osterreith Complex Figure Test (ROCFT)

Kasai et al. (2006) compared performance on the ROCFT (Corwin & Bylsma, 1993, as cited in Lezak et al., 2004) in elderly healthy and mildly cognitively impaired (MCI) Japanese participants. They found that MCI subjects, who in general have a high probability of later developing AD (Petersen et al., 1999), performed more poorly than healthy participants. They thus concluded that the ROCFT is a useful test for detecting impaired learning and memory for visual material in early AD. There are no published research studies documenting the performance of South Africans diagnosed with AD on the ROCFT.

Cross Cultural Effects on Digit Span

Ostrosky-Solis and Lozano (2006), who tested 2574 Mexican subjects aged between 16 and 96 years, found that level of literacy and number of years of education were significantly positively correlated with performance on the digit span subtest of the Wechsler Adult Intelligence Scale – Third Revision (WAIS-III; Wechsler, 1997). They also compared the results of their study to results from North America, South America, Europe, Australasia and South Africa, and found

that there were also differences in performance between cultures, even when controlling for age and level of education. These data suggest that the digit span test is not culture-free, as is commonly assumed (Ostrosky-Solis & Oberg, 2006).

Cross Cultural Effects on the Trail Making Test (TMT)

Even supposedly non-verbal neuropsychological tests are influenced by culture. Recent normative studies of the TMT (Strauss, Sherman & Spreen, 2006) have been performed for Korean, Japanese and Australian elderly populations. These studies have noted the important positive effect of educational level on performance, especially on part B of the test (Hashimoto et al., 2006; Hester, Kinsella, Ong, & McGregor, 2005; Seo et al., 2006). The studies noted that, although Trails A was still useful for poorly educated older adults, Trails B was often not completed by those participants. These findings have particular relevance for the application of the TMT to the South African context, given that many patients receiving outpatient care at public hospital memory clinics are poorly educated and of low socioeconomic status. There are no published research studies documenting the performance of South Africans diagnosed with AD on the TMT.

Neuropsychological Testing in a Multi-Cultural context

In taking the multiple factors that affect neuropsychological test performance into account, researchers have adopted two main approaches: (i) creating cross-culturally fair tests, or identifying and applying among existing tests only those that are cross-culturally fair, or (ii) obtaining new norms, taking into account multiple demographic and cultural variables (Lezak et al., 2004; Nell, 2000).

The first approach has relevance when using commonly used Western tests to assess “rural and illiterate or semi-literate” and extremely non-Westernized populations, such as rural African tribes or Amazonian Indians (Ardila, 1996; Shuttleworth-Jordan, 1996, p. 96). Shuttleworth-Jordan (1996), however, cautions against the assumption that commonly used Western tests need to be completely replaced by newly developed culturally fair tests. She points out that developing new tests requires enormous time and resources, and that cultural effects are

diminished due to the post-apartheid acculturation process occurring in South Africa, and to increasing urbanization and globalization.

Thus, the second main approach, of obtaining new norms, seems more workable. Clinicians in developing countries are under increasing pressure to provide neuropsychological assessments of brain-damaged patients for health, rehabilitation and legal reasons (Nell, 2000). Without new culturally fair tests, the most efficient solution would be to develop new culturally-appropriate norms for existing well-standardized and validated tests. Indeed, many authors have voiced the urgent need for new norms (Ardila, 1996; Boone et al., 2007; Nell, 2000; Shuttleworth-Jordan, 1996).

SPECIFIC OBJECTIVES AND HYPOTHESES

Although some non-Western normative studies have been done on older adults, few studies have been conducted using South African populations, where there are both white Western and African-based cultures. Thus, there is considerable scope for future research in this area.

The Albertina and Walter Sisulu Institute of Ageing in Africa (IAA), part of the Division of Geriatric Medicine at the University of Cape Town, conducts a weekly outpatient Memory Clinic at Groote Schuur hospital. This public hospital services patients mainly from the Western Cape region of South Africa. Patients presenting at the IAA Memory Clinic are mostly over the age of 50, and are of different races, cultures, socio-economic status, education level, and languages. Because of a variety of memory complaints, patients are referred to the Clinic either by community health practitioners or by their family and primary caregivers. A standard battery of neuropsychological tests is administered as part of a broader assessment by a multidisciplinary team of geriatricians, psychiatrists, psychologists and other healthcare professionals. The aim of the neuropsychological screening is to determine the pattern of cognitive impairment and differentiate between possible diagnoses, including AD and vascular dementias.

This study describes how healthy and demented South African older adults perform on the particular battery of neuropsychological tests used at the IAA Memory Clinic. As such, it has a

dual purpose. Firstly, it is the start of a much larger and more comprehensive project, providing a first step toward collecting normative data for older South African adults on some of the most commonly-used memory clinic neuropsychological tests. Secondly, it aims to answer the question of whether these tests are useful in helping to detect early AD in the South African population. We focus specifically on AD because internationally it is the most prevalent of the dementias, and this is reflected in the patient population we see at the IAA memory clinic (Zillmer, Spiers & Culbertson, 2001).

We tested the following set of hypotheses: (1) Healthy young adults will show better performance, in general, on the Memory Clinic neuropsychological test battery than will healthy older adults, (2) healthy older adults will show better performance, in general, on the Memory Clinic battery than will AD patients, (3) demographic variables relevant for the South African context, such as age, sex, race, SES, education level, quality of education, and home language, will together have a combined effect, apart from main group differences, on test performance in general, (4) the diagnostic efficiency of the Memory Clinic battery will be greater when calculated using South African cutoffs (derived from data obtained in this study) compared with International cutoffs.

DESIGN & METHODOLOGY

Design

To determine the usefulness of the Memory Clinic battery in the South African population, it was necessary to describe the differences in test performance that occur due to normal ageing, as well as differences that occur due to the effects of AD. This study therefore had a quasi-experimental design, comparing the performance of three groups on the Memory Clinic battery of neuropsychological tests. These groups were (1) healthy young adults (YA), (2) healthy older adults (OA), and (3) older adults previously and independently diagnosed with possible or probable Alzheimer's disease (AD). Each group was administered the same battery of neuropsychological tests under standardized conditions. It is important to note, however, that neuropsychological test data was obtained retrospectively from the IAA database. We investigated the effect on performances due to socio-cultural factors, such as education,

language, race and socioeconomic status. Finally, we investigated the diagnostic efficiency of the particular tests in the battery, specific to identifying AD. This involved testing the sensitivity, specificity, positive predictive power and negative predictive power of each of the tests, using South African cut-off scores (obtained from our healthy older adult means and standard deviations; Kessel & Zimmerman, 1993; Strauss et al., 2006).

Participants

Young adult participants (YA; n = 29) were UCT psychology undergraduate students, recruited via the Department of Psychology's Student Research Participation Programme (SRPP). Healthy older adult participants (OA; n = 25) were functionally independent and community-dwelling individuals recruited from a range of local communities in the Western Cape region of South Africa. In all locations, personal contacts were utilized to encourage suitable individuals to participate. AD participants (n = 24) had been assessed at the IAA Memory Clinic between January 2005 and July 2007, and had an independent diagnosis of either possible or probable AD. Mild, moderate, and severe AD cases were included in the sample. Neuropsychological data for these participants was therefore obtained from the Memory Clinic database. Table 1 presents the demographic characteristics of all three groups.

Exclusion Criteria

MMSE. Any participants from the healthy groups (YA and OA) who obtained an MMSE score of below the cutoff for cognitive impairment of 24/30, were excluded (Lezak et al., 2004). This may, however, be an unfair cut-off for some, due to the effects of age, gender, education, race/ethnicity, language and socioeconomic status (Black et al., 1999; Lezak et al., 2004). There are also difficulties when the MMSE is translated or used in cultures different from the original North American population (Gibbons et al., 2002). To be on the safe side, however, the established cut-off was used to exclude participants who may be cognitively impaired from our healthy adults groups to ensure the integrity of the normative data collected in this study (Lezak et al., 2004). Two potential OA participants were excluded on the basis of this criterion.

Depression. Depression can negatively affect performance on neuropsychological testing, especially with regard to recall memory tests (Hertell, 2000, as cited in Lezak et al., 2004). For

this reason, moderately to severely depressed (ie. those with BDI-II scores of more than 20) participants were screened out of the YA and OA groups. Two potential YA participants, and one potential OA participant were excluded on the basis of this criterion.

Other. Participants with a history of multiple head injuries, memory and learning difficulties, neurological disorders, and/or psychological disorders were also be excluded. Three potential OA participants, and two potential YA participants were exclude on the basis of these criteria.

Materials

Demographic Questionnaire. A demographic questionnaire was used to obtain information about a participant's age, sex, race, handedness, highest level of education completed, education history, home language, socioeconomic status, and health history. Socioeconomic status was estimated by the total monthly income of the participant's household (low: R0 - 999; medium: R1000 - 5000; high: R5000 +). Quality of education for the YA participants was estimated from the type of educational institutions (public = low; private = high), regardless of race. However, quality of education was estimated for the OA and AD participants based on race, and so it was estimated that white participants had a high quality of education and non-white participants had a low quality of education. This estimation was based on the segregated schooling system of apartheid in South Africa, which provided higher quality of education for whites compared to non-whites.

Beck Depression Inventory-II (BDI-II). The BDI-II (Beck, Steer, & Brown, 1996) is a 21-item scale used widely in clinical settings to assess depression. Moderate depression has been shown to influence neuropsychological test performance (Lezak, 2004). Lezak et al. (2004) report test-retest reliability coefficients for this scale between 0.74 and 0.93. Validity coefficients range from 0.81 to 0.51 in clinical studies. The BDI-II also correlates with the first version of the BDI at 0.93. It is used widely in clinical settings in South Africa, and a number of studies using the inventory have been done on in the South African population (e.g., Faure & Loxton, 2003; Pillay, 2001).

Mini-Mental State Examination (MMSE). The MMSE (Folstein et al., 1975) is a briefly administered screening test for basic mental status. It assesses orientation in time and space, as well as basic memory functions such as immediate and delayed recall, attention, language, calculation, and constructional abilities (Lezak et al., 2004). It uses standardized administration instructions and simple standardized scoring (Knight, McMahon, Green, & Murray-Skeaff, 2006). According to Lezak et al. (2004), performance is more or less stable in healthy adults, but deteriorates over time in AD patients.

Lezak et al. (2004) report test-retest reliability coefficients across different studies of 0.83 to 0.89 in the original sample used in standardization. The MMSE also correlates highly with Verbal and Performance IQ scores on the WAIS (Verbal IQ and MMSE, $r = 0.776$; Performance IQ and MMSE, $r = 0.660$; Folstein et al., 1975). Although the MMSE is used widely in South African clinical settings, no published research studies using South African samples were found.

CLOX Task. In this test, which assesses both visuospatial ability and executive control function, individuals are first asked to draw a clock (CLOX1), and are then required to copy a clock drawn by the examiner (CLOX2; Royall et al., 1998). Instructions are standardized, and the two drawings are scored on the same 15-point system. Lezak et al. (2004) report that (a) performance on this task is consistently worse in AD compared with healthy older adults, and (b) AD patients perform significantly better on CLOX2 compared with CLOX1.

Royall (1998) reports high internal consistency ($r = 0.82$), and high between-rater reliability (CLOX1, $r = 0.94$; CLOX2, $r = 0.93$). CLOX correlates highly with the MMSE and another test of cognitive severity ($r = 0.82$-0.85) (Royall, 1998). Once again, while the CLOX is used widely in South African clinical settings, no published research studies using South African samples were found.

Verbal Fluency. Tests of verbal fluency in memory clinic settings usually consist of a semantic fluency test, where the individual is instructed, for instance, to name as many animals as they can in 1 minute; and a phonemic fluency test, where, for instance, they are instructed to name as many words starting with the letter ‘F’ in 1 minute. These tests assess the “speed and ease of

verbal production" as well as the way individuals organize their thinking (Lezak et al., 2004, p. 518).

Duff-Canning, Leach, Stuss, Ngo, and Black (2004) found that Canadian AD patients performed significantly worse on both semantic and phonemic fluency tests than did normal controls. They also found that semantic fluency was significantly worse than phonemic fluency in AD patients, even for mildly demented subjects with MMSE scores greater than 24.

Mitrushina et al. (2005) report high inter-rater reliability ($r = 0.98$), high test-retest reliability (r ranged from 0.72 to 0.74 across different studies) and high internal consistency ($r = 0.83$) for various verbal fluency tests. Tests of verbal fluency are used widely in clinical settings in South Africa. Only one published research study was found that used verbal fluency tests in South African English-Zulu bilingual participants between the ages of 18 and 45 (Bethlehem, de Picciotto, & Watt, 2003). They found that the semantic verbal fluency was a successful test for screening brain injury in individuals with a high proficiency in English.

Rey Auditory Verbal Learning Test (AVLT). This measure assesses verbal learning and memory. It involves learning and free recall over five trials of a 15-word list, a sixth trial of this list after interference, and a 30-minute delay followed by free recall and recognition trials of the first list (Estevez-Gonzalez, Kulisevsky, Boltes, Otermin, & Garcia-Sanchez, 2003; Knight et al., 2006; Schmidt, 1996, as cited in Lezak, et al., 2004).

Lezak et al. (2004) report high test-retest reliability across different studies ($r = 0.61$-0.86 for part one; $r = 0.51$-0.72 for part two). They also report that the AVLT correlates highly with other measures of learning, including the California Verbal Learning Test (r ranges from 0.3 to 0.47 across different studies). The AVLT is used widely in South African clinical settings. Cassimjee, Couzens, Smith, and Wagner (2004) used the AVLT successfully to assess South African coronary artery bypass patients after surgery.

Rey-Osterreith Complex Figure Test (ROCFT). This test involving copying, immediate recall and 30-minute delayed recall of a complex geometric figure assesses perceptual organization and

visual memory (Corwin & Bylsma, 1993, as cited in Lezak, et al., 2004). Since the IAA memory clinic battery does not include the ROCFT, or any test of visual memory, we included the ROCFT in our battery and administered it to our healthy young and older adults. This was to begin the process of collecting South African norms so that this test can be used at the IAA memory clinic in future, since it has been found to be useful in detecting AD in clinical settings (Kasai et al., 2006; Mitrushina et al., 2005). The E. M. Taylor (1959, as cited in Mitrushina, et al., 2005) scoring method was used.

The inter-rater reliability for the Taylor scoring criteria is also high (*r* ranges from 0.80 to 0.99 across different studies; Mitrushina et al., 2005). Test-retest reliability is, however, moderately low (Mitrushina et al., 2005). The ROCFT is used widely in South African clinical settings, but no published studies of this population were found.

Digit Span. In this study, we used a version of the digit span forward and backward tests that form part of the Wechsler intelligence and memory batteries. These tests assess immediate recall, working memory, and attention (Ostrosky-Solis & Lozano, 2006; Wechsler, 1997). In the forward digit span test, examinees are required to verbally repeat a series of digits, with increasingly long digit strings. In the backward digit span test, examinees are required to verbally repeat the series of digits in the reverse order, again, with increasingly long digit strings.

Lezak et al. (2004) report test-retest reliability coefficients for the forward digit span of between 0.66 and 0.89. The South African version of Wechsler Adult Intelligence Scale (SA-WAIS; Human Sciences Research Council, 1969) includes digit span tests, and is administered widely in South African clinical settings (Shuttleworth-Jordan, 1995).

Trail-Making Test (TMT). Part A of this test (Trails A) requires the examinee to draw lines between numbers, in sequence, on a page as quickly as he/she can. Part B of the test (Trails B) requires the examinee to switch between connecting numbers and letters in series. It assesses visual search, motor coordination, and attention (Knight et al., 2006; Spreen & Strauss, 1998).

Lezak et al. (2004) report high reliability coefficients across different studies for the TMT (r = 0.60-0.90). The TMT is used widely in South African clinical settings. Cassimjee et al. (2004) used the TMT successfully to assess South African coronary artery bypass patients after surgery.

Procedure

Each participant was individually tested in a psychology laboratory room or in a similarly quiet room in their place of residence. We ensured that the room was similar to the one in which the Memory Clinic patients were tested, and so, it was quiet, uncluttered and contained a desk and two chairs. The researcher and participant sat opposite each other with a desk between them. The researcher briefly described the study and then gave the participant the consent form to read through and sign. The participant then completed the demographic questionnaire, the BDI-II, and the MMSE, and then the testing began. Testing took approximately 90 minutes to complete. The battery was administered and scored according to standard procedures for each individual test outlined in Lezak et al. (2004) and/or in the individual test manuals. The order of test administration is detailed in Table 2. After testing was completed, the researcher explained the aims and rationale of the study, and gave the participant an opportunity to ask questions.

RESULTS

Data Analysis.

Group comparisons analysis involved simple one-way between-group ANOVAs on each of the neuropsychological test outcome variables. Homogeneity of variance was violated for many of the tests, but we carried on with the analysis, because ANOVA is relatively robust to this kind of violation. Where there was a significant overall effect between groups, planned comparisons (using Tukey's HSD) were done to evaluate the relationships between YA and OA groups, as well as OA and AD groups, as hypothesized. In order to reduce the family wise error rate, a Bonferroni correction of $\alpha/4=0.0125$ was used.

To test the overall effect of a group of demographic variables (age, sex, race, SES, education quality, education level ad home language) on test performance, across the groups, I created a general linear model. Each of the demographic variables was added to the model, as a predictor,

in the above order, and the main effects were observed, by adding the sums of squares of each of the variables, in order to account for the amount of variance in test scores explained by these demographic variables.

To determine how useful individual tests are at detecting AD in a South African context, we used our data to determine sensitivity, specificity, positive predictive power and negative predictive power for each of the tests (Kessel & Zimmerman, 1993). Sensitivity, is defined as "the percentage of ill persons who are identified by the test as ill," while specificity is defined as "the percentage of non-ill persons correctly identified by the test as non-ill" (Kessel & Zimmerman, 1993, p.395). Positive predictive power, is defined as "the percentage of individuals classified by the test as ill who truly are ill", and negative predictive power is defined as "the percentage of individuals classified by the test as non-ill who truly are non-ill" (Kessel & Zimmerman, 1993, p.395). Cut-offs for the individual tests were calculated by subtracting 1.5 standard deviations from the OA group mean for each individual test. Any participant who obtained a score lower than the cutoff was labeled as "positive," in terms of a diagnosis of AD using only that individual test; those who obtained a score higher than the cutoff were labeled "negative".

There was no data missing for the YA group. One of the participants in the OA group completed only the CLOX and none of the other neuropsychological tests. Another of the participants in the OA group did not complete the AVLT, but completed all other tests. The true number of AD participants who completed the each of the tests is given in Table 3. Missing data for the AD group was due to the participant being unable to complete the test. When a test was discontinued, the scores were excluded from the analysis. However, when a participant had started the TMT parts A and B, but was unable to complete the test, a maximum score of 300 seconds was given. No participants were excluded as outliers. Data analysis was performed using the Statistica Version 7.0 package.

Descriptive Statistics.

Descriptive statistics were obtained for each of the groups on each of the test variables, except for the ROCFT, and are reported in Table 3. Table 4 reports descriptive statistics on the ROCFT, for the YA and OA groups only. The general trend, upon inspecting the means for each group,

was that the participants in the YA group performed better than the participants in the OA group, who in turn, performed better than participants in the AD group. On the ROCFT, participants in the YA group once again obtained higher scores than did participants in the OA group.

Overall Group Comparisons

The results of these analyses are reported in Tables 5 and 6. Omnibus *F* tests were significant for all the dependent variables, except for CLOX difference score. Furthermore, there were no significant differences between YA and OA group means, and OA and AD group means, on the CLOX difference score. Effect sizes were mostly very large, ranging from 15.82 % (RAVLT-LOT, $p < 0.0011$) to 87.39% (RAVLT-DRI, $p < 0.0001$). It is also interesting to note that the greatest effect was found on the RAVLT variables (except for LOT, which may be problematic by definition, and will be discussed later).

YA and OA Group Comparisons

CLOX. I predicted that participants in the YA group would perform significantly better on both CLOX1 and CLOX2, and would have a lower CLOX difference score, than participants in the OA group. However, as shown in Table 5, there were no significant YA-OA between-group differences on any of the CLOX measures.

Verbal Fluency. I predicted that participants in the YA group would perform significantly better on both letter and category fluency than participants in the OA group. However, as shown in Table 5, there was no significant YA-OA between-group difference on letter fluency. On the other hand, YA participants performed significantly better than OA participants on category fluency.

Digit Span. I predicted that participants in the YA group would perform significantly better than participants in the OA group on both the forward and backward digit span tests. However, as shown in Table 5, there were no significant YA-OA between-group differences on forward or backward digit span.

TMT. I predicted that participants in the YA group would perform significantly better than participants in the OA group on both Trails A and Trails B. However, as shown in Table 5, while there was no significant YA-OA between-group difference on Trails A, participants in the YA group performed significantly better than participants in the OA group on Trails B.

AVLT. I predicted that participants in the YA group would perform significantly better on all the AVLT test variables, than participants in the OA group. However, as shown in Table 5, participants in the YA group performed significantly better on RAVLT total words, STPR, LTPR, DRI and adjusted recognition, than participants in the OA group. There was no significant YA-OA between-group difference on RAVLT-LOT.

OA and AD Group Comparisons.

CLOX. I predicted that OA participants would perform better than AD participants on the CLOX1, that there would be no difference in performance between OA participants and AD participants on CLOX2, and that OA participants would have a lower CLOX difference score compared to AD participants. The first prediction was confirmed; however, as shown in Table 5, participants in the OA group performed significantly better than participants in the AD group on CLOX2, and, there were no significant OA-AD between-group differences on CLOX difference.

Verbal Fluency. I predicted that OA participants would perform better than AD participants on letter and category fluency. This prediction was confirmed.

Digit Span. I predicted that there would be no significant OA-AD between-group differences on both forward and backward digit span. The prediction was confirmed for forward digit span only; on the backward digit span test, OA participants performed significantly better than AD participants.

TMT. I predicted that there would be no significant OA-AD between-group differences on Trails A, and that OA participants would perform significantly better than AD participants on Trails B. However, as shown in Table 5, OA participants performed significantly better than AD participants on both Trails A and B.

AVLT. I predicted that participants in the OA group would perform significantly better than participants in the AD group on all the AVLT test variables. This prediction was confirmed for all of these variables except RAVLT-LOT, where there were no significant between-group differences.

Effect Size of Demographic Variables on Test Performance

The effect sizes (using adjusted R^2 as an estimate of ω^2) for each of the tests are reported in Table 7. Figure 1 shows the effect of group as a predictor for test performance compared with the overall effect of the demographic variables as predictors. In general, the effect of the group is larger than the effect of the demographic variables. However, Figure 1 demonstrates that the overall effect of demographic variables on test performance is also quite substantial. For example, in RAVLT total words, group status alone explains 85 percent of the variance in test performance, whereas the demographic variables together account for 74 percent of the variance. Category fluency, Trails B, and the AVLT were the most highly affected, having demographic variables accounting for more than 40 percent of the variance. Furthermore, on CLOX2, demographic variables seem to account for more of the variance in performance than group status does.

Diagnostic Efficiency

The diagnostic efficiency results are given in Table 8. Specificity was generally low (< 0.85), except for on RAVLT-DRI and Trails A. On the other hand, sensitivity was high (> 0.85) for all of the test variables. In general, both positive and negative predictive power was high (all > 0.77). Positive predictive power was particularly low for the RAVLT-LOT (0.50), however.

DISCUSSION

Status of the Hypotheses

YA and OA test performance. We have seen that, in general, the means for the YA group were greater than the OA group, although the YA-OA between-group differences on approximately

half of the test variables were not statistically significant. YA participants performed significantly better on category fluency, Trails B, and the RAVLT variables. One might thus conclude that these tests are the most sensitive to normal ageing processes, such as slowed information processing speed (Zhang, Davis, Salthouse & Tucker-Drob, 2007).

OA and AD test performance. We have seen that, in general, the means for the OA group were greater than those for AD group, and that the OA-AD between-group differences in performance on the majority of tests were highly statistically significant. There was no significant difference found between OA participants and AD participants on the forward digit span test, although this was in accordance with what was predicted: This test measures short-term memory, which would be only be affected in more severe cases of AD (Ostrosky-Solis & Lozano, 2006; Wechsler, 1997). Interestingly, though, OA participants performed better than AD participants on the backward digit span test, which may be because it requires more complex mental activity (i.e., working memory), involving "mental tracking" (Lezak, 2004, p.358). Overall, we may conclude that the Memory Clinic test battery is sensitive to differences in test performance due to AD, apart from differences due to age.

Problematic test variables. There were no significant between-group differences between the groups on CLOX difference scores. The RAVLT-LOT measure seemed to be problematic in general, and did not yield any highly significant results. This lack of statistical significance may be explained by the definition of this measure. Participants who displayed consistently good performance on trials I to V would have a lower LOT than participants who performed more poorly in general. LOT may, therefore, only be useful for measuring the learning ability for those already identified as cognitively impaired.

Effect of demographic variables on test performance. It has been widely shown in the cross-cultural neuropsychology literature that many individual biological and socio-cultural variables impact on test performance. We were interested in the general effect of a group of variables, specific to the South African context, on test performance. We were interested to observe the effect of the group of variables together, since many of these variables are interrelated in the real clinical population. For instance, taking South African apartheid history into account, race may

be a proxy for other socio-cultural variables such as SES and education. While it may be useful to separate out the effects of the individual variables, it seems to be a more workable and clinically applicable approach to deal with these variables together. It was found, in general, that age, sex, race, SES, education quality, education level and home language together have a significant impact on test performance. This may suggest that demographic variables may contribute to a diagnosis of AD in the South African population. We did not investigate the direction of the overall effect of these variables, and so cannot conclude whether they are likely to lead to over-pathologizing or under-pathologizing. We may assume, however, from the general trend in non-Western cross-cultural studies reviewed above, that older individuals with non-white race, low SES, low education level and quality, and non-English home language would perform poorer on the Memory Clinic battery. Demographic variables specific to the South African context, particularly race, SES, education and language, will need to be taken into account for the diagnosis of AD using neuropsychological test performance. Studies of similar magnitude as the MOANS need to be carried out in South Africa to clarify and quantify the impacts of these variables (and including IQ) on neuropsychological test performance.

Diagnostic efficiency using South African norms. We have shown that the South African norms obtained from our healthy OA participants had high sensitivity, high positive and negative predictive power, and low specificity. It is important to consider the South African context in which tests are administered. Patients who eventually present at the IAA Memory Clinic are often disadvantaged, have usually been on a waiting list for months, and may have traveled long distances to be assessed by highly qualified professionals for a variety of memory complaints. In this case, it would be more beneficial to assess individuals using a more sensitive measure than a more specific measure, because those identified by the neuropsychological test battery as "positive" for possible AD could then receive further assessments provided by the government health services, which could, over time, confirm or disconfirm the diagnosis. If a more specific (but less sensitive) measure was used, patients who really do have AD may be identified as "negative" for possible AD, in other words as "healthy", and thereafter no longer seek further treatment. We may therefore conclude that South African norms are more appropriate for use in the IAA Memory Clinic, and in the South African context in general, than Western norms.

Status of the Major Aims

One of the major aims of this study was to provide a first step toward collecting normative data for older South African adults on some of the most commonly-used memory clinic neuropsychological tests. South African norms (or cutoffs) for the tests in the IAA Memory Clinic battery were calculated and tabulated in Table 8; we showed that these norms were useful for identifying AD in the South African population. Additionally, South African norms for the ROCFT are provided in the form of means and standard deviations (see Table 4), which may be a useful reference should the ROCFT be included in the Memory Clinic battery in the future. In order to obtain more accurate and appropriate norms, to make more accurate diagnoses, a much larger sample of South African older adults needs to be assessed.

Future cross-cultural research should focus on collecting more data from the South African population, in order to provide more powerful norms. By keeping cell sizes equal and having a larger sample, more powerful conclusions could be drawn from between-group comparisons data. Familywise error rate may have been high, with a total number of four statistical tests done on each test variable. Although a Bonferroni correction was made, so that tests were evaluated at the 0.0125 significance level, the risk of a Type I error may also have been high because all the test variables were derived from performance by related to the same participants. This was not such a serious problem for this study, however, because analysis is still in the explorative stage. It was therefore more beneficial to do as many tests as possible rather than restrict the analysis to only the most statistically elegant procedures.

Another aim of the study was to determine whether the tests in the IAA Memory Clinic battery were useful for detecting AD in the South African population. We saw that, in general, YA, OA and AD participants performed as predicted on the majority of the tests. Indeed, diagnostic efficiencies of these tests were high when using South African norms. We can therefore conclude that the test battery can be successfully used for identifying AD in the South African context, and that by developing South African norms these tests may be even more effective.

REFERENCES

Ardila, A. (1996). Towards a cross-cultural neuropsychology. *Journal of Social and Evolutionary Systems, 19,* 237-248.

Beck, A.T., Steer, R. A. & Brown, G. K. (1996). *BDI –II manual*. San Antonio: Psychological Corporation.

Bethlehem, D., de Picciotto, J., & Watt, N. (2003). Assessment of verbal fluency in bilingual Zulu-English speakers. *South African Journal of Psychology, 33,* 236-240.

Black, S. A., Espino, D. V., Mahurin, R., Lichtenstein, M. J., Hazuka, H. P., Fabrizio, D., et al. (1999). The influence of non-cognitive factors on the mini-mental state examination in older Mexican-Americans: findings from the Hispanic EPESE. *Journal of Clinical Epidemiology, 52,* 1095-1102.

Boone, K. B., Victor, T. L., Wen, J., Razani, J., & Ponton, M. (2007). The association between neuropsychological scores and ethnicity, language, and acculturation variables in a large patient population. *Archives of Clinical Neuropsychology, 22,* 355-365.

Cassimjee, N., Couzens, C. L., Smith, F. J., & Wagner, C. (2004). Neuropsychological outcomes of coronary artery bypass grafting. *Health SA Gesondheid, 9*, 3-14.

Duff-Canning, S. J., Leach, L., Stuss, D., Ngo, L., & Black, S. E. (2004). Diagnostic utility of abbreviated fluency measures in Alzheimer disease and vascular dementia. *Neurology, 62*, 556-562.

Estevez-Gonzalez, A., Kulisevsky, J., Boltes, A., Otermin, P., & Garcia-Sanchez, C. (2003). Rey verbal learning test is a useful tool for differential dignosis in the pre-clinical phase of Alzheimer's disease: Comparison with mild cognitive impairment and normal ageing. *International Journal of Geriatric Psychiatry, 18*, 1021-1028.

Faure, S., & Loxton, H. (2003). Anxiety, depression and self-efficacy levels of women undergoing first trimester abortion. *South African Journal of Psychology, 33*, 28-38.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189 – 198.

Gaither, S. (Ed.) (2007). *Clinical diagnosis and management of Alzheimer's Disease* (3rd ed.). Hampshire: Thomson Publishing Services.

Gibbons, L. E., van Belle, G., Yang, M., Gill, C., Brayne, C., Huppert, F. A., et al. (2002). Cross-cultural comparisons of the Mini-Mental State Examination in United Kingdom and United States participants with Alzheimer's disease. *International Journal of Geriatric Psychiatry, 17,* 723-728.

Hashimoto, R., Meguro, K., Lee, E., Kasai, M., Ishii, H., & Yamaguchi, S. (2006). Effect of age and education on the Trail Making Test and determination of normative data for Japanese elderly people: the Tajiri project. *Psychiatry and Clinical Neurosciences, 60*, 422-428.

Hester, R. L., Kinsella, G. J., Ong, B., & McGregor, J. (2005). Demographic influences on baseline and derived scores from the Trail Making Test in healthy older Australian adults. *The Clinical Neuropsychologist, 19*, 45-54.

Hodges, J. R. (Ed.) (2001). *Early-onset dementia: A multidisciplinary approach.* New York: Oxford University Press.

Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2000). *Principles of neural science* (4th ed.). New York: McGraw-Hill.

Kasai, M., Meguro, K., Hashimoto, R., Ishizaki, J., Yamadori, A., & Mori, E. (2006). Non-verbal learning is impaired in very mild Alzheimer's disease (CDR 0.5): normative data from the learning version of the Rey-Osterrieth Complex Figure test. *Psychiatry and Clinical Neurosciences, 60*, 139-146.

Kessel, J. B. & Zimmerman, M. (1993). Reporting errors in studies of the diagnostic performance of self-administered questionnaires: Extent of the problem, recommendations for standardized presentation of results, and implications for the peer review process. *Psychological Assessment, 5,* 395-399.

Knight, R. G., McMahon, J., Green, T. J., & Murray-Skeaff, C. (2006). Regression equations for predicting scores of persons over 65 on the Rey Auditory Verbal Learning Test, the Mini-Mental State Examination, the Trail Making Test and semantic fluency measures. *British Journal of Clinical Psychology, 45*, 393-402.

Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological assessment* (4th ed.). New York: Oxford.

Manly, J. J. (2005). Advantages and disadvantages of separate norms for African Americans. *The Clinical Neuropsychologist, 19*, 270-275.

Mitrushina, M., Boone, K. B., Razani, J., & D'Elia, L. F. (2005). *Handbook of normative data for neuropsychological assessment* (2nd ed.). New York: Oxford University Press.

Nell, V. (2000). *Cross-cultural neuropsychological assessment: Theory and practice.* New Jersey: Lawrence Erlbaum Associates.

Oberg, G., & Ramirez, M. (2006). Cross-linguistic meta-analysis of phonological fluency: normal performance across cultures. *International Journal of Psychology, 41*, 342-347.

Ostrosky-Solis, F., & Lozano, A. (2006). Digit span: effect of education and culture. *International Journal of Psychology, 41*, 333-341.

Ostrosky-Solis, F., & Oberg, G. (2006). Neuropsychological functions across the world-common and different features: from digit span to moral judgment. *International Journal of Psychology, 41,* 321-323.

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangolos, E. G., Kokman, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology, 56,* 303-308.

Pillay, A. L. (2001). Psychological symptoms in recently diagnosed cancer patients. *South African Journal of Psychology, 31*, 14-18.

Roselli, M., & Ardila, A. (2003). The impact of culture and education on non-verbal neuropsychological measurements: A critical review. *Brain and Cognition, 52,* 326-333.

Royall, D. R., Cordes, J. A., & Polk, M. (1998). CLOX: An executive clock drawing task. *Journal of Neurology, Neurosurgery and Psychiatry, 64,* 588-594.

Royall, D. R., Espino, D. V., Polk, M. J., Verdeja, R., Vale, S., Gonzales, H., et al. (2003). Validation of a Spanish translation of the CLOX for use in Hispanic samples: the Hispanic EPESE study. *International Journal of Geriatric Psychiatry, 18*, 135-141.

South African Wechsler Adult Intelligence Scale (SAWAIS) Manual. (1969). Johannesburg: National Institute for Personnel Research, Human Sciences Research Council.

Seo, E. H., Lee, D. Y., Kim, K. W., Lee, J. H., Jhoo, J. H., Youn, J. C., et al. (2006). A normative study of the Trail Making Test in Korean elders. *International Journal of Geriatric Psychiatry, 21*, 844-852.

Shuttleworth-Edwards, A. B., Kemp, R. D., Rust, A. L., Muirhead, J. G. L., Hartman, N. P., & Radloff, S. E. (2004). Cross-cultural effects on IQ test performance: A review and preliminary normative indications on WAIS-III test performance. *Journal of Clinical and Experimental Neuropsychology, 26,* 903-920.

Shuttleworth-Jordan, A. B. (1995). On not reinventing the wheel: A clinical perspective on culturally relevant test usage in South Africa. *South African Journal of Psychology, 26,* 96-111.

Spreen, O., & Strauss, E. (1998). *A compendium of neuropsychological tests* (2nd ed.). New York: Oxford University Press.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests* (3rd ed.). New York: Oxford University Press.

Steinberg, B., & Bieliauskas, L. (2005). IQ-based MOANS norms for multiple neuropsychological instruments. *The Clinical Neuropsychologist, 19*, 277-279.

Steinberg, B., Bieliauskas, L., Smith, G. E., & Ivnik, R. J. (2005). Mayo's older Americans normative studies: Age- and IQ-adjusted norms for the trail-making test, the stroop test, and MAE controlled oral word association test. *The Clinical Neuropsychologist, 19*, 329-377.

Steinberg, B., Bieliauskas, L., Smith, G. E., Ivnik, R. J., & Malec, J. F. (2005). Mayo's older Americans normative studies: Age- and IQ-adjusted norms for the auditory verbal learning test and the visual spatial learning test. *The Clinical Neuropsychologist, 19*, 464-523.

Steinberg, B., Bieliauskas, L., Smith, G. E., Langelloti, C., & Ivnik, R. J. (2005). Mayo's older Americans normative studies: Age- and IQ-adjusted norms for the Boston naming test, the MAE token test, and the judgment of line orientation test. *The Clinical Neuropsychologist, 19,* 280-328.

Wechsler, D. A. (1997). *Wechsler Adult Intelligence Scale-III.* San Antonio: The Psychological Corporation.

Wechsler, D. A. (1981). *Wechsler Adult Intelligence Scale - Revised.* New York: Psychological Corporation.

Zhang, Z. Z., Davis, H. P., Salthouse, T. A., & Tucker-Drob, E. M. (2007). Correlates of individual, and age-related, differences in short-term learning. *Learning and Individual Differences, 17,* 231–240.

Zillmer, E. A., Spiers, M. V., & Culbertson, W. C. (2001). *Principles of neuropsychology.* (2nd ed.) U.S.A.: Thomson Wadsworth.

NOTES

1. Steinberg, Bieliauskas, Smith, Langelloti, et al. (2005) describe how measurements of educational level and quality can be problematic as a predictors of neuropsychological performance. For individuals with the same level of education in years, there may be variations in academic achievement, quality of education received, marking stringency, and curriculum contents. An individual may not have reached a higher level of education because they were unable to cope academically, or, other factors may have prevented them from doing so (such as poverty, or Apartheid laws preventing Blacks from attending university).

Table 1.

Sample Demographic Characteristics

Demographic Variable		YA (n=29)	OA (n=25)	AD (n=24)
Age				
	Range	18 – 24	65 - 86	69-85
	Mean (SD)	20 (1.28)	74.68 (5.05)	75.7(4.87)
Sex				
	Male	15	9	10
	Female	14	16	14
Race				
	White	16	12	4
	Non-White	13	13	20
SES				
	High	20	8	8
	Medium	7	9	11
	Low	2	8	5
Education quality				
	High	9	2	0
	Low	20	23	24
Education mean years (SD)		13.28 (1.25)	11.04 (2.09)	10.42 (2.96)
Home Language				
	English	23	18	19
	Other	6	7	5

Table 2.

Order of Test Administration

TEST	Test Variables measured
1. CLOX	
CLOX 1	CLOX 1
CLOX 2	CLOX 2
	CLOX Difference: (CLOX2 – CLOX1)
2. Verbal fluency (letter "F")	No. of words
3. Semantic fluency (animals)	No. of words
4. Rey Auditory Verbal Learning Test (AVLT) – I	
Trials I to V	Total words
List B	Learning over trials (LOT)
Trial VI	Short Term Percentage Recall (STPR)
	Long Term Percentage Recall (LTPR)
	Delayed Recall Index
	Adjusted Recognition
5. Digit Span	
Forward	No. of digits
Backward	No. of digits
6. Trail Making Test (TMT)	
Part A	Time taken (s) – part A
Part B	Time taken (s) – part B
7. AVLT - II	
Trial VII	
Recognition	

Table 3.

Descriptive Statistics

Test	YA Mean	(SD)	n	OA Mean	(SD)	n	AD Mean	(SD)	n
CLOX1	12.379	(1.8011)	29	11.8	(2.1409)	25	7.9565	(4.7623)	23
CLOX2	14.4828	(0.6877)	29	13.16	(1.1431)	25	10.8261	(4.5392)	23
CLOX difference[a]	2.1034	(1.9702)	29	1.3600	(1.8903)	25	2.6400	(3.3526)	23
Letter fluency	13.3103	(4.9360)	29	12.9167	(5.2826)	24	5.9575	(3.8399)	16
Category fluency	23.0690	(4.5271)	29	17.2083	(5.4929)	24	6.7917	(3.4513)	24
Forward digit span	6.6207	(0.6769)	29	6.0	(0.9325)	24	5.4	(1.3139)	20
Backward digit span	5.2069	(1.0816)	29	4.4167	(1.4116)	24	3.1000	(1.5526)	20
Trails A – time taken	30.0	(10.8858)	29	42.1667	(9.1731)	24	142.35	(97.6386)	20
Trails B – time taken	58.5517	(14.1537)	29	125.2917	(46.2643)	24	259.6000	(60.8652)	20
RAVLT – Total words[b]	60.4138	(7.6696)	29	34.7391	(6.4893)	23	18.4	(8.2104)	20
RAVLT – LOT[c]	16.2759	(8.9038)	29	13.4348	(4.5809)	23	8.2105	(6.1335)	19
RAVLT – STPR[d]	93.8897	(13.0842)	29	59.4043	(23.6662)	23	24.3333	(31.5235)	21
RAVLT – LTPR[e]	98.1103	(13.3960)	29	60.0348	(22.5255)	23	19.500	(19.7337)	20
RAVLT – DRI[f]	26.3103	(3.8555)	29	10.5652	(4.5809)	23	1.8571	(3.0048)	21
RAVLT – Adj recognition[g]	14.0345	(1.7624)	29	5.9565	(3.1835)	23	-5.4211	(7.8550)	19

Note.

a. 'CLOX difference' is calculated by subtracting the raw score of CLOX1 from CLOX2.

b. 'Total words' is the sum of the words recalled over trials I to V.

c. 'Learning Over Trials' is calculated as follows: (Sum of words from trials I to V) - (5x words recalled on trial I).

d. 'Short Term Percentage Recall' is the percentage words recalled on trial VI compared with trial V.

e. 'Long Term Percentage Recall' is the percentage words recalled on trial VII compared with trial V.

f. 'Delayed Recall Index' is the sum of the short term (trial VI) and long term (trial VII) delayed recall trials.

g. 'Adjusted Recognition' calculated by subtracting false positives from true positives.

Table 4.

Descriptive Statistics for ROCFT

Test	YA (n=29)		OA (n=23)	
	Mean	(SD)	Mean	(SD)
ROCFT Copy - time	144.97	(62.55)	186.96	(59.97)
ROCFT Copy - score	34.72	(1.39)	29.89	(6.34)
ROCFT Immed. Recall - time	138.66	(55.04)	134.35	(79.44)
ROCFT Immed. Recall - score	25.6	(4.44)	13.11	(5.19)
ROCFT Delayed Recall - time	101.93	(42.86)	104.91	(60.88)
ROCFT Delayed Recall - score	25.1	(5.69)	12.82	(4.91)

Table 5.

Results of Simple Analysis of Variance

	Levene's test for homogeneity of variance		Omnibus F test				Tukey's Planned Comparisons		
TEST	F	p	F	p	df	Effect size (Adj.R^2)	Prediction:	p	df
CLOX1	19.9076	<0.0001*	14.9588	<0.0001*	2,74	0.2690	YA>OA	<0.7700	74
							OA>AD	<0.0002*	74
CLOX2	20.2800	<0.0001*	12.8540	<0.0001*	2,74	0.2380	YA>OA	<0.1552	74
							OA=AD	<0.0074*	74
CLOX Difference	7.4105	<0.0012*	2.2948	<0.1126	2,74	0.0318	YA<OA	<0.5136	74
							OA<AD	<0.0926	74
Letter fluency	1.2800	<0.2850	13.5017	<0.0001*	2,66	0.2710	YA>OA	<0.9533	66
							OA>AD	<0.0002*	66
Category fluency	1.9934	<0.1435	84.4025	<0.0001*	2,74	0.6870	YA>OA	<0.0001*	74
							OA>AD	<0.0001*	74
Forward digit span	7.576	<0.0011*	9.5350	<0.0002*	2,70	0.1917	YA>OA	<0.0592	70
							OA=AD	<0.1088	70
Backward digit span	0.4654	<0.6300	14.8061	<0.0001*	2,70	0.2772	YA>OA	<0.0877	70
							OA=AD	<0.0049*	70
Trails A – time taken	55.0203	<0.0001*	31.4901	<0.0001*	2,70	0.4586	YA<OA	<0.6707	70
							OA=AD	<0.0001*	70
Trails B – time taken	21.4621	<0.0001*	134.7810	<0.0001*	2,70	0.7880	YA<OA	<0.0001*	70
							OA<AD	<0.0001*	70

Note.
*Bonferroni correction: Since four statistics were calculated for each of the individual neuropsychological test variables, $\alpha/4=0.0125$, and therefore, for significance $p< 0.0125$.

Table 6.

Results of Simple Analysis of Variance: continued

	Levene's test for homogeneity of variance		Omnibus F test				Tukey's Planned Comparisons		
TEST	F	p	F	p	df	Effect size (Adj. R^2)	Prediction:	p	df
RAVLT – Total words	0.3477	<0.7075	197.2070	<0.0001*	2,69	0.8468	YA>OA	<0.0001*	69
							OA>AD	<0.0001*	69
RAVLT – LOT	7.5756	<0.0011*	7.5786	<0.0011*	2,68	0.1582	YA>OA	<0.3226	68
							OA>AD	<0.0499	68
RAVLT – STPR	12.9205	<0.0010*	56.1716	<0.0001*	2,70	0.6051	YA>OA	<0.0001*	70
							OA>AD	<0.0001*	70
RAVLT – LTPR	8.1790	<0.0006*	128.4307	<0.0001*	2,62	0.7821	YA>OA	<0.0001*	69
							OA>AD	<0.0001*	69
RAVLT – DRI	2.6250	<0.0797	246.9789	<0.0001*	2,69	0.8739	YA>OA	<0.0001*	69
							OA>AD	<0.0001*	69
RAVLT – Adjusted recognition	24.1533	<0.0001*	103.7855	<0.0001*	2,62	0.7505	YA>OA	<0.0001*	69
							OA>AD	<0.0001*	69

Note.
*Bonferroni correction: Since four statistics were calculated for each of the individual neuropsychological test variables, $\alpha/4=0.0125$, and therefore, for significance $p< 0.0125$.

Table 7.

General Effect of Demographic Variables[a] on Test Performance

Test	Effect size (Adj. R^2)	F	p	df
CLOX1	0.1504	2.41670	<0.0256*	8,56
CLOX2	0.3893	6.09914	<0.0001*	8,56
CLOX Difference	-0.0265	0.79345	<0.6106	8,56
Letter fluency	-0.0045	0.96417	<0.4731	8,56
Category fluency	0.4373	7.21675	<0.0001*	8,56
Forward digit span	0.1978	2.97279	<0.0076*	8,56
Backward digit span	0.1789	2.74350	<0.0125*	8,56
Trails A – time taken	0.1960	2.95049	<0.0080*	8,56
Trails B – time taken	0.4766	8.28375	<0.0001*	8,56
RAVLT- Total words	0.7410	23.88621	<0.0001*	*8,56*
RAVLT- LOT	0.0537	1.45437	<0.1950	*8,56*
RAVLT- STPR	0.4870	8.59329	<0.0001*	*8,56*
RAVLT- LTPR	0.5380	10.31540	<0.0001*	*8,56*
RAVLT – DRI	0.7621	26.63124	<0.0001*	*8,56*
RAVLT- Adj. Recognition	*0.4862*	*8.56911*	*<0.0001**	*8,56*

Note.
a. Demographic variables included: age, sex, race, SES, education quality, education level, home language.
*Bonferroni correction: Since four statistics were calculated for each of the individual neuropsychological test variables, $\alpha/4=0.0125$, and therefore, for significance $p< 0.0125$.

Table 8.

Diagnostic efficiency of tests using South African cutoff scores

	SA cut-off	Specificity	Sensitivity	Positive predictive power	Negative predictive power
CLOX1	8.589	0.522	0.944*	0.800	0.823
CLOX2	11.445	0.348	0.981*	0.889*	0.779
Letter fluency	4.993	0.438	0.981*	0.875*	0.855*
Category fluency	8.969	0.667	0.963*	0.889*	0.867*
RAVLT – Total words	25.005	0.800	0.981*	0.941*	0.930*
RAVLT – LOT	6.564	0.421	0.852*	0.500	0.807
RAVLT – STPR	23.905	0.571	0.981*	0.923*	0.855*
RAVLT – LTPR	26.247	0.750	0.981*	0.938*	0.914*
RAVLT – DRI	3.694	0.857*	1.000*	1.000*	0.947*
RVLT – Adjusted Recognition	1.181	0.778	0.981*	0.933*	0.930*
Forward digit span	4.601	0.550	0.963*	0.846	0.852*
Backward digit span	2.299	0.750	0.981*	0.938*	0.914*
Trails A – time taken	55.926	0.850*	0.925*	0.810	0.942*
Trails B – time taken	194.688	0.750	0.981*	0.938*	0.914*

Note.
*Scores > or = 0.85 indicate high diagnostic efficiency.

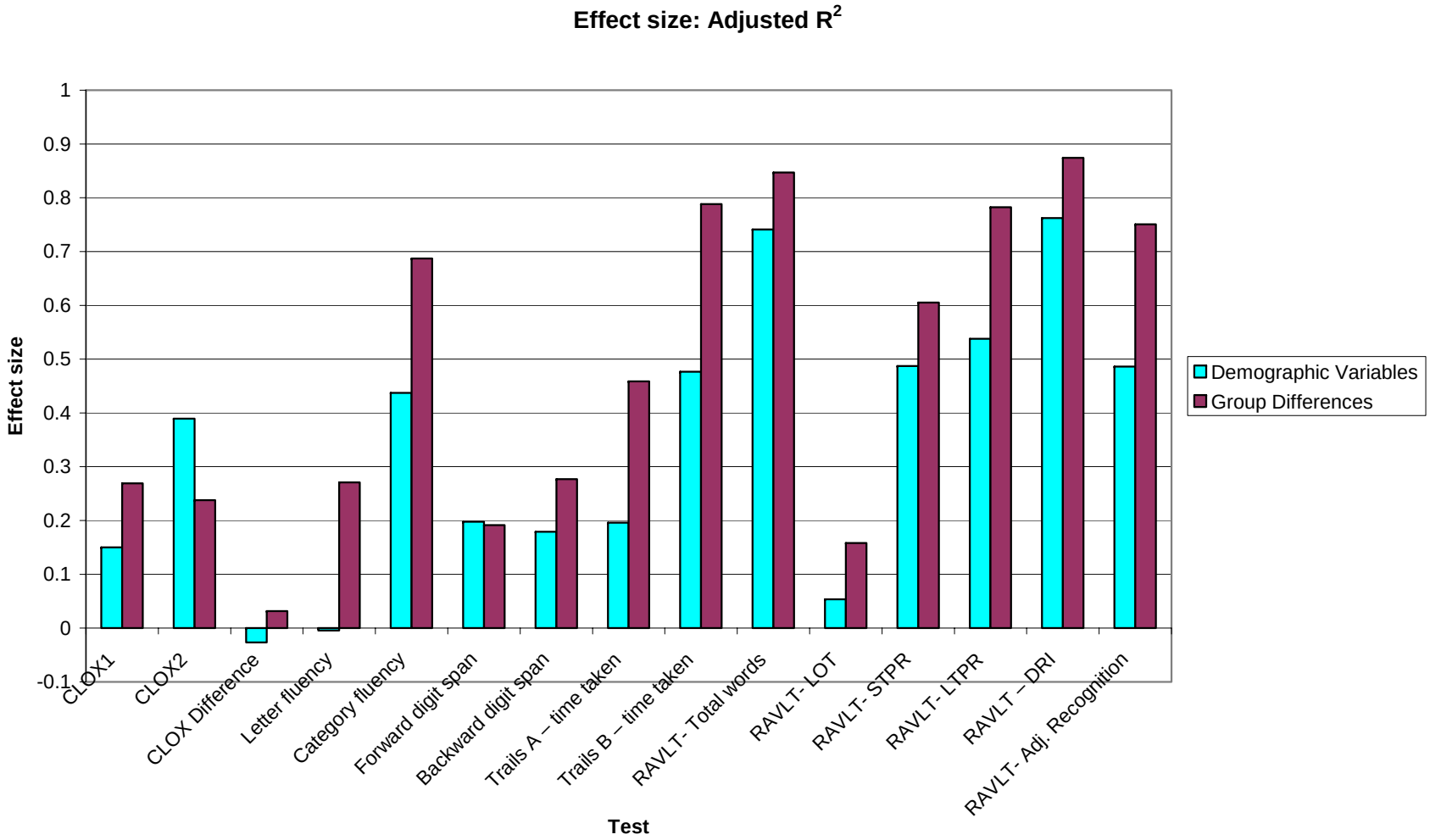

Figure 1. Effect sizes of Group Compared with Demographic Variables

Plagiarism Declaration

1. I know that plagiarism is wrong. Plagiarism is using another's work and to pretend that is one's own.

2. I have used the American Psychological Association as the convention for citation and referencing. Each significant contribution to, and quotation in, this report from the work, or works of other people has been attributed and cited and referenced.

3. This report is my own work.

4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

Signature: _____

Date: _____