

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

19

**DEVELOPMENT AND VALIDATION OF A
NEUROCOGNITIVE SCREENING BATTERY:
LANGUAGE AND SPATIAL COGNITION**

Jill Mosdell

MSDJIL001

A minor dissertation submitted in partial fulfilment of the requirements for the award of
the degree of Master of Arts in Psychological Research

Faculty of Humanities
Department of Psychology
University of Cape Town

2008

Declaration:

This work has not been submitted in whole, or in part, for the award of any degree. It is my own work. Each significant contribution to, and quotation in, this dissertation from the work, or works, of other people has been attributed, and has been cited and referenced.

Signature:

Date:

Contents:

1.) Abstract	Pg 3
2.) Introduction and Literature Review	Pg 4
3.) Aim	Pg 33
4.) Methodology	Pg 35
5.) Results	Pg 54
6.) Discussion	Pg 83
7.) Conclusion	Pg 107
8.) References	Pg 108
9.) Appendices	Pg 122

1.) Abstract

There is great need for an effective neurocognitive screening tool in South Africa. The outdated and diagnostically limited Mini Mental State Exam (MMSE) is currently the only widespread form of cognitive screening. This and virtually all other neurocognitive tests widely used in South Africa were developed in Europe and North America and are inappropriate for South Africa's population. The present study aided in the development and validation of the first theory-driven neurocognitive screening tool specifically designed for South Africans, focusing on the Language and Spatial Cognition sections of the battery. Two tests — the Naming Test and the 3-D Analysis Test — were both adapted and re-piloted before their insertion into the Neurocognitive Battery ahead of its validation.

The re-piloting sample comprised equal numbers of isiXhosa-, English- and Afrikaans-speaking controls (n=30). Results showed that the two tests did not show language bias. However, the 3-D Analysis Test was further modified as many controls performed poorly on it. Test-retest and inter-rater reliability were shown for both the Language section and the Spatial Cognition section. Validation was established by showing that the Language and Spatial Cognition sections could discriminate between controls and between patients with left versus right-sided lesions (n=45). This study provides strong initial support for the reliability and validity of the Language and Spatial Cognition sections of the Neurocognitive Battery, while offering pertinent information and suggestions for the further development of this screening tool.

Development and Validation of a Neurocognitive Battery: Language and Spatial
Cognition

2.) Introduction and Literature review

Need for neuropsychological assessment in South Africa

At Groote Schuur Hospital, the Neuropsychology Division of the Neurology Department is referred patients with a wide spectrum of pathologies, all in need of neuropsychological assessment. Such patients can range from those with strokes, seen at the hospital's Stroke Unit, to patients suffering from a variety of dementias (e.g. Parkinson's disease, Alzheimer's disease, HIV/AIDS dementia etc). Epidemiologically speaking, South Africa has high occurrences of illnesses that present with neurocognitive deficits. For example, a recent study estimates that sixty people die from cerebrovascular accidents (CVAs) every day in South Africa (Heart and Stroke Foundation South Africa, 2007). The need for neuropsychological assessment is not limited to neurological patients; referrals from neurosurgical colleagues (most often head trauma) and psychiatry are also common.

Motor vehicle accidents are also highly prevalent in South Africa, with many of the people involved in such accidents sustaining head trauma. In 2003, it was estimated that 25 000 fatal car injuries occurred (National Injury Mortality Surveillance System Annual Report, 2003). Unfortunately, there are no statistics on the number of South Africans involved in car accidents; however, almost 75% of worldwide road accidents occur in

developing countries. Vehicle accidents are also the main cause of death in adolescents and adults (Odero, Garner & Zwi, 1997). Alarming, in South Africa, a quarter of all buses and minibuses are involved in vehicle collisions every year (Council for Scientific and Industrial Research, 1995 in Mba, 2004).

South Africa also has the ignominy of having one of the highest rates of HIV infection anywhere in the world. The prevalence rates range from 11.4% to 27.9% (South African Department of Health, 2003; Nelson Mandela Study, 2002 from <http://www.avert.org/safricastats.htm>). Ten% to 20% of HIV-infected individuals and 30% to 40% of AIDS patients develop neurological disorders and cognitive dysfunction (Mochan, Modi & Modi, 2003). In 2003, it was reported that approximately 2 million adults in South Africa were infected with tuberculosis and HIV.

It is therefore vital that our healthcare institutions — especially a large public hospital such as Groote Schuur — have access to assessment tools that adequately assess neurocognitive function in the South African population. Such tools are needed by a large number of patients. Neuropsychological testing/screening is imperative in helping patients by tracking the nature and possible advancement of their disease and in helping medical professionals select the most appropriate therapy and/or medication. For example, neuropsychological performances are essential in predicting post-stroke functioning (Barker-Collo & Feigin, 2006). Sound neurocognitive assessments are crucial in helping the patient's family and loved ones to understand the nature of his or her

cognitive dysfunction and can therefore aid in the development of compensation and management methods.

Limitations of Mini-Mental State Examination (MMSE)

Before the commencement of the overarching study to which this current study contributes, Groote Schuur Hospital, along with other South African public hospitals, used the MMSE as their only form of cognitive screening and assessment. The MMSE is the most popular cognitive assessment tool of neurocognition in the world (Lezak, 1995). Nysa, et al. (2005) point out that cognitive disorders are important indicators of long-term outcome in stroke patients. They say that the MMSE was developed as a screening tool for dementia and is now widely used as a general screening tool for all cognitive deficits. However, they question its validity as a general cognitive screening tool in neurological patients and psychiatric patients (Faustman, Moses, & Csemansky, 1990; Grace et al., 1995). Nysa et al. (2005) argue that the MMSE is unable to discriminate between cognitively impaired and intact patients suffering from acute stroke. The MMSE has also been reported to be insensitive to right-hemisphere lesions (Grace et al., 1995).

Research further suggests that the MMSE is appreciably influenced by ethnicity and education (Stewart et al. 2002). Literature on the efficacy of the MMSE used in South Africa is lacking; however international research has shown that older patients (Bleecker, Bolla-Wilson, Kawas, & Agnew, 1988 in Lezak, 1995) and patients with low levels of education attain lower scores on the MMSE (Ostrosky-Solis, Lopez-Arango & Ardila, 2000). Further studies have shown that African Americans and females are more likely to

score false positives than Caucasian-American and male populations (Anthony et al., 1982 in Lezak, 1995).

Contrasting MMSE results across populations should therefore be approached tentatively, especially regarding factors such as ethnicity. Watts (2005) maintains that there is a dire need for specific South African assessment and intervention procedures in our country. She points out that clinicians have resorted to using existing tests and the norms that have been developed in the West. This may result in misdiagnoses being made.

A further notable flaw of the MMSE being used as the sole screening tool is that the examination is not driven by any underlying neurocognitive theory. Wind, et al. (1997) argue that the MMSE is quite limiting in being able to diagnose dementia since it does not look at each cognitive subtype individually, but rather only provides one overall score.

Neurologists at Groote Schuur Hospital have found the MMSE to be inadequate as a sole neurocognitive screening instrument and therefore asked their Neuropsychology Division to create a more appropriate screening tool. A screening instrument was needed that could sufficiently and meaningfully assess neurocognitive functioning in the South African context while at the same time be theoretically sound.

Initial Pilot Study

This current study is a continuation of an initial phase (completed two years ago) of the overarching study. The overall aim of the larger ongoing study was twofold: firstly, to create culturally fair neurocognitive tests for use in a neurocognitive screening battery for South Africa; and secondly, to validate this battery in its entirety. The goal of the initial phase was to adapt certain neuropsychological tests in a neurocognitive battery — at the time known as the Groote Schuur Neurocognitive Battery — and to adapt these tests to be appropriate culturally and linguistically for the South African context. Tests of both language and spatial cognition were adapted to eliminate bias and to become more appropriate for a South African context.

The cultural appropriateness of the tests in the battery was questioned from the beginning on the basis of the practicing neuropsychologists' clinical experiences. Hence, the tests were adapted with the help of neuropsychologists from the Neuropsychology Division at Groote Schuur as well as isiXhosa- and Afrikaans-language and culture experts from the South African Languages Department at UCT.

As part of the larger study, the language and cultural experts were also consulted in order to translate and back-translate the entire Neurocognitive Battery into isiXhosa and Afrikaans because it is crucial that the battery be available in the preferred language of our patients. Swartz, Drennan and Crawford (1997) and Swartz (1998) caution the incredible influence language has on assessing and treating patients. A patient being assessed in a language s/he is not comfortable communicating in could lead to poor

performances on tests. This would not be due to cognitive deficits but to defective translation or a misunderstanding of instructions. This is especially important when adapting and translating tests that assess language.

Fortuny et al. (2005) argue that people often take translations for granted as it is expected that test construction is done by individuals who are educated in the local dialects and that the translations will be proofread repeatedly. They recommend that in order to ensure proper translation takes place, experts in the local language must be part of test construction, translation and back-translation of the tests; specialists in terminology and test construction must be a part of the process and proofreading needs to be done by educated, native speakers. Swartz (1998) argues that everyone involved in the translation process must be sensitive to translation errors and cultural effects. These recommendations directed the translation process by ensuring that the neuropsychologists, language and cultural experts worked together in the translation and back-translation process. The language and cultural experts at UCT were also native speakers of the languages they taught.

After translation was completed, the newly adapted and translated tests were piloted on thirty neurologically healthy controls as part of the previous study completed two years ago. Overall, the results at the time showed controls performing normally, unlike their performances on the original neuropsychological tests, where many control subjects fared extremely poorly. Two neuropsychological tests, however, were still shown to have some cultural/educational bias, namely, the 3-D Analysis Test and the Naming Test. Controls

in the study scored an average of 17.3 from 30 (normal isiXhosa participants scored an average of 13.1) for the original Boston Naming Test. The new Naming test's scores were more promising, with controls scoring an average of 27, and all three language groups (English, Afrikaans and isiXhosa) scored similarly with a standard deviation of 1.85.

However, a few of the new items were deemed inappropriate. For example, isiXhosa controls had problems identifying the item 'hippopotamus', while Afrikaans and English controls did not. The recommendations from the pilot study were that these items be re-worked. The problems with the new 3-D Analysis Test were determined to be the small size of the blocks and the distortion in perspective of some of the items. The recommendations that followed were to enlarge and simplify the block design, as many controls had difficulties passing the test. It is important that these tests be further adapted before validation is attempted.

Effects of culture and language on neuropsychological tests

The patient population seen at Groote Schuur is multicultural, with patients coming from both urban and rural areas. The adaptations of the Naming and 3-D Analysis tests should consider language and cultural influences. Harris (1983) in Rosselli and Ardila (2003) explains culture as the way people live as a group, which involves values, attitudes, behaviours and approaches to thinking. Reynolds (2000) in Rosselli and Ardila (2003) argues that new neuropsychological tests need to be created for specific cultures because

group differences do not denote a difference in ability but are rather indicative of the bias of the tests.

Ardila (2003) in Rosselli and Ardila (2003) maintains that the differences between cultural groups on non-verbal tests are the result of a multifaceted interaction between cultural factors and brain organisation. The concept of culture is not easily defined and is entangled with socio-economic status and ethnicity (Gasquoine, 1999). Agranovich and Puente (2007) argue that cultural factors influence one's performances on neuropsychological assessments.

Rossellia and Ardila (2003) argue that a culture-free test is an illusion. They are critical of previous studies that have tried to achieve culture-free tests by focusing on non-verbal tests such as visuo-spatial tests. They argue that copying figures and other visuo-spatial tests are school-dependant abilities. To illustrate, Henry (2001) argues that there are cultural differences regarding Rey Figure performances.

Brauer Boonea, Victor, Wen, Razani and Pont'on (2007) report that there is a blatant effect of culture/ethnicity on test performances where African-Americans perform less well than Caucasians on a range of cognitive tests (e.g. Boston Naming Test) (Whitfield et al., 2003 in Brauer Boonea et al., 2007). They report that level of education is one of the strongest influences on the test performance. Brauer Boonea et al. (2007) show that African-American, Hispanics and Asians score 9-10 points lower on the Boston Naming Test after levels of education and age are adjusted for. Australian populations also

perform worse on the Boston Naming Test than North American and European populations (Cruice, Worrall & Hickson, 2000).

Pahl and Kara (1992) in Carter et al. (2005) investigated the suitability of the Renfrew Word Finding Scale (RWFS) in South Africa. Over half of the normal participants scored low enough to suggest vocabulary deficits. Pahl and Kara (1992) in Carter et al. (2005) also found that the Indian children from the sample scored much lower than the White children even though they came from the same language and socio-economic background. This illustrates how enormously culture can affect neuropsychological performance.

Carter et al. (2005) offer suggestions in order to provide culturally valid tools. They mention the necessity of having mother-tongue speakers of the intended population to help in the development of the tests. These mother-tongue speakers must also be familiar with the culture of the population that the tests are intended for. They recommend local teachers. They further argue that the adapted tests then need to be piloted on a representative sample of the intended population.

There is a dire need for modifications of neuropsychological tests specific to culture and geographic location (Kohnert, Hernandez & Bates, 1998). Therefore, Western-based tests need to be adapted for cultural factors (Wong & Fujii, 2004). However, language is just one element of culture. Formal education is another crucial factor that greatly influences cognitive performance (Rosselli & Ardila, 2003).

Effects of education on neuropsychological tests

Most patients at Groote Schuur Hospital have low levels of education, and often even those with higher levels have been subjected to a low quality of education due to the former government's policies. There is a strong link between neuropsychological performance and level of education (Ostrosky, Ardila, Rosselli, pez-Arango & Uriel-Mendoza, 1998; Byrd, Sanchez & Manly, 2005) and quality of education (Agranovich & Puente, 2007). It has been found that education levels affect performances on the Boston Naming Test (Talberg, 2005) and language tests in general (Ostrosky et al., 1998). Mansur et al. (2005) mention that visual confrontation naming tests (such as the Boston Naming Test) are especially sensitive to demographic variables such as education level.

In studies of African Americans, it was shown that performances were influenced by education and cultural experience (Kennepohl, Shore, Nabors & Hanks, 2004; Manly, Byrd, Touradji & Stern, 2004). The norms for the Boston Naming Test were developed from a North American population for two education levels: those with less than twelve years of education and those with more than twelve years of education. It is thus important to adapt tests in order to limit the influence low levels of education would have on tests, such as showing false positives (Killgore & Adams, 1999).

Levels of education have also been argued to influence factors such as constructional ability, language and calculation abilities (e.g., Ardila, Rosselli, & Rosas, 1989; Rosselli et al., 1990; Lecours et al., 1987a, 1987b, 1988 in Ostrosky-Solis, Ardila, Rosselli, Lopez-Arango & Uriel-Mendoza, 1998). As Ostrosky-Solis et al. (1998) caution, if

educational factors are not taken into account, brain pathology might be diagnosed in normal groups with low education levels.

Comparison of neuropsychological performances to inappropriate norms

There are no norms for the different cultural groups in South Africa and therefore no scores to compare patients by gender, ethnicity, education level or even first language. Manly and Echemendia (2007) argue that a current problem is the lack of empirical research regarding ethnic minorities in neuropsychological assessment. Cognitive measures have not been validated for most ethnic minorities. The adapted tests (as with most neuropsychological tests) were constructed from a Western, white, middle-class perspective with their norms determined from North American populations (Gasquoin, 1999; Worrall, Yiu, Hickson & Barnett, 1995 in Storms, Saerens and De Deyn, 2004). It is dangerous to compare South African performances on the tests to their North American norms. Heaton, Ryan, Grant and Matthews (1996) in Heaton, Avitable, Grant and Matthews (1999) argue that the application of norms based on one specific ethnic group to another one can lead to results falsely suggesting pathology.

Rey, Feldman, Rivas-Vazquez, Levin and Benton (1999) argue that considering cultural factors with regards to neuropsychological assessment is paramount, as medical and legal decisions are regularly based on the results of the assessment. They argue that in the United States, neuropsychological tests are often adapted and translated for a Hispanic population without any validation. They also mention that Hispanic patients' performances are commonly compared to norms derived from English-speaking

Americans. They disapprove of the practice of researchers comparing Hispanic patients' performances on translated batteries with the norms derived from English-speaking Americans. There is no evidence to support the claim that translations of psychometric tests are comparable to the original English version (Rey, Feldman, Rivas-Vazquez, Levin & Benton, 1999).

Giannakou and Kosmidis (2006) state that in Greece, tests developed in the United States are used on Greek populations. Their performances are compared inappropriately to unsuitable norms, which leads to misdiagnoses among the Greek population. They argue that cultural factors need to be addressed before neuropsychological tests can be used. They further argue that nonverbal tests are very sensitive to cultural factors (Ardila & Moreno, 2001; Jacobs et al., 1997; Rosselli & Ardila, 2003 in Giannakou & Kosmidis, 2006).

A great deal of research in cross-cultural neuropsychology has been based on a Caucasian-comparative approach where ethnic groups' neuropsychological performances are compared to Caucasian group performances (Byrd, Sanchez & Manly, 2005; Henry, 2001). The common result is that these ethnic groups perform more poorly on the tests than their Caucasian counterparts.

South African norms should therefore be determined, or preferably, these tests should be modified for specific cultures (Brauer Boonea et al., 2007; Mansur, Radanovic, Taquemori, Greco & Araujo, 2005). Adaptation and re-piloting is, however, not the only

goal of this study. The Neurocognitive Battery comprising the adapted tests needs to be shown to possess validity and reliability — i.e. whether the battery with its constellation of tests is in fact able to identify/localise lesion sites based on the neurocognitive deficits seen with left versus right hemisphere lesions, and whether it is able to do so consistently.

Validity for a neurocognitive battery

Russell, Russell and Hill (2005) define a battery differently to a group of tests. They argue that if one uses two or more tests to reach an interpretative understanding, one is using a battery. The strengths of a battery are many: it provides one with an opportunity to understand patterns in brain functioning from the relationships between the tests. The authors argue that even if individual tests have been validated, if these tests are used together to form a battery, then validation must be shown for the battery (Russell, Russell & Hill, 2005).

Russell, Russell and Hill (2005) discuss the double dissociation method where, for example, the results of at least two tests are applied to two hemispheres of the brain with the purpose of ascertaining the relationship between the hemispheres and the tests (Teuber, 1955, 1975 in Russell, Russell & Hill, 2005). Russell et al. (2005) argue that in the 1950s, a comparison of tests was used to localise lesions in the brain. This is known as Reitan's rule; that is, a minimum of two tests are used with a constant relationship with each other in order to localise lesions accurately.

Christensen and Caetano (1999) and Russell et al. (2005) contend that this method is successful because the different tests used have different sensitivities to varying brain abnormalities and different areas of the brain. They argue that while a specific brain lesion will affect the tests to some degree, it will affect the test that is most sensitive to that area of the brain. The goal is to compare patients' neurocognitive functioning to their own weaknesses and strengths (Christensen & Caetano, 1999).

Therefore, they argue that it is not the final score in a test battery that is important for a diagnosis but the ratio of scores. They thus conclude that a precise approach to localising lesions is to compare tests that are sensitive to a known location with tests that are insensitive to that brain location. Russell, Russell and Hill (2005) conclude that neuropsychological batteries need to show validity in order to be construed as dependable.

Problems establishing validity and reliability in neuropsychological tests

There are significant problems with establishing validity and reliability in neuropsychological tests and screening batteries. For example, patients' symptoms and performances on tests due to neuropsychological abnormalities can vary day to day, thus affecting the supposed reliability of tests. Kaszniak (1989) in Lezak (1995) argues that the customary requisites of a test to show validity and reliability and have suitable norms are difficult to attain in neuropsychological assessments. Lezak (1995) further argues that invaluable assessment techniques that have been able to pick up on neuropsychological

deficits have been developed from clinical experiences. However, these techniques have not been standardised.

Many “good” tests that show validity and reliability may not provide meaningful insight into neuropsychological functioning: “Test batteries that generate summed or averaged scores based on a clutch of discrete tests provide another example of good reliability (the more scores, the more reliable their sum) of a score that conveys no neuropsychological relevant information unless the score is either so low or so high that the level of the contributing scores is obvious” (Lezak, 1988b in Lezak, 1995:120). Luria (1999) argues that the reason for neuropsychological examinations is to describe the neuropsychological impairments in patients and to be able to *understand* the underlying reasons for the impairment. Hebben and Milberg (2002) state that many neuropsychologists are not convinced that the use of normative data composed in a battery is the best way to detect brain damage and that many psychometric tests are not derived from a theory regarding brain and behaviour.

Guedalia, Finkelstein, Drukker and Frishberg (2000) argue that reliability is best shown with patients with focal neuropsychological dysfunction in neuropsychological batteries (Kane, 1991; Reitan & Wolfson, 1985 in Guedalia et al., 2000). Guedalia et al. (2000) argue that neuropsychological batteries that are purely psychometric may be ineffective in discerning focal brain problems that are coupled with diffuse or serious brain pathologies. A qualitative approach is important in being able to form hypotheses related to the behavioural components of the brain pathology (Guedalia et al. 2000).

Luria's approach

For this study, Luria's hypothetico-deductive approach to neuropsychological assessment will be adopted, which is agreed by many to be the most detailed (Glozman, 1999). Luria's influence in neuropsychology and neuropsychological assessment is widely acknowledged (Kotik-Friedgut, 2006). The goal of the assessment is to identify observable, specific abnormalities in brain functioning and to link them with their neuro-anatomical correlates. For example, an abnormality with speech production is associated with Broca's area in the brain. Therefore, the patients' performances on tests must be able to elicit functioning of a multitude of different neuropsychological skills. A qualitative approach to assessment must be employed to ensure that a dysfunction is exactly identified and its precise severity noted. This leads to a description and elucidation of the patient's individual abnormalities.

Luria provides a theory of brain organisation that maintains that complex behaviours are comprised of more basic functions (Hebben & Milberg, 2002). Luria's approach consisted of a series of observations ordered into decision trees. These decision trees echo the organisation of brain function (Hebben & Milberg, 2002). The goal is to see whether the patient could perform the tests given to him or her and then to eliminate the different possible reasons for the dysfunction observed. Such an approach is adopted by the present research, which comprises of decision trees that represent various brain functions.

The assessment should infer where in the brain the lesion that caused this neurocognitive deficit would be. Tupper (1999) argues that Lurian neuropsychological assessment is

theory-driven meaning that the assessment is designed in accordance with a theoretical understanding of how mental functions are organised in the brain. It aims to link the neuropsychological assessment with the functional system of the brain. Lurian assessment is also oriented around single cases. The psychometric approach, on the other hand, uses purely quantitative measures and mainly uses group studies.

Glozman (1999) argues that Luria's neuropsychological assessment is thought to be the most flexible and comprehensive form of neuropsychological assessment. Glozman (1999) suggests that a Lurian approach assumes that a qualitative analysis needs to be made regarding the patient's performance. "Luria's approach and methods presuppose a qualitative analysis of the symptom under study, that is, the analysis and comparison of primary disorders – those immediately connected with the impaired factor – with secondary disorders, or those that emerge in accordance with the systemic organization of higher mental functions" (Christensen, 1974; Homskaya, 1987; Luria, 1965, 1966/1980, 1973b in Glozman, 1999: 23).

The two most popular quantitative fixed batteries are the Halstead-Reitan Neuropsychological Battery and the Luria-Nebraska Neuropsychological Battery (Hebben & Milberg, 2002). According to Glozman (1999), attempts were made to quantify and standardise Luria's methods by developing the Luria-Nebraska Neuropsychological Battery (LNNB). The thinking was that this was the only way Lurian methods would gain approval from Western neuropsychologists (Glozman, 1999).

The main disadvantage of this battery is that its tests are often condensed into a pass-or-fail score that disallows the qualitative analysis that typifies Luria's assessment methods (Glozman, 1999). However, Glozman (1999) argues that quantitative analysis of neuropsychological functioning is important because it is able to ascertain the change in neurocognitive functioning of the patient for follow-up or after rehabilitation. She argues that a combination of qualitative and quantitative methods in neuropsychological assessment is needed to keep up with the dynamic aspects of neuropsychology. One of the ways to do so is to apply Luria's qualitative procedures to psychometric measures.

The LNI (Luria's Neuropsychological Investigation) is a qualitative neuropsychological assessment that uses the hypothesis-testing approach (Christensen & Caetano, 1999). The LNI's advantage is that it is based on theoretical principles of neurocognitive functioning (Christensen & Caetano, 1999). The LNI is flexible, brief to administer and economical. However, the mentioned disadvantages of the qualitative LNI are that scoring is subjective and no validation studies have been done. Christensen and Caetano (1999) argue that qualitative approaches are not less scientific than quantitative methods. Luria always selected tests that could be easily done by people without brain pathology; and that allowed component analysis on the tests' performances.

Luria (1966, 1973) in Guedalia et al. (2000) argues that a patient failing a test does not necessarily mean that he or she has a brain lesion that corresponds to the neuropsychological test. Most tests require the use of a number of brain activities. The way that a patient attempts the test gives a lot of information to a number of cognitive

functions. The specific areas of brain lesions are identified by looking at the unity of the cognitive deficits on a number of tests.

Korkman (1999) mentions Luria's view on higher mental activity as dynamic functional organisations that echo focal brain activity. Korkman (1999) elucidates that the "brain is viewed as a functional mosaic the parts of which in various combinations provide the neural basis of cognitive processes (Luria, 1963, pp ix-x; 1973, pp. 11, 26-30 in Korkman 1999: 90). Korkman (1999) notes that the way a patient fails a test is an important indicator of his or her brain pathology. Neuropsychological assessments are also supposed to be useful in determining the corresponding areas of the brain that are affected. Korkman (1999) argues that brain imaging can be used nowadays to verify the conclusions of a neuropsychological assessment. For example, a diagnosis of aphasia with comprehension problems can be verified with a lesion found in Wernicke's area. Korkman (1999) has developed a Lurian-inspired assessment battery using a pass/fail decision assessment.

Tupper (1999) argues that Luria's work has been approached more than any other to provide a metatheoretical understanding to neuropsychology. He further adds that no other neuropsychologist has "a better claim about the respectability, usefulness, and longevity of his or her writings" (Tupper, 1999: 2). Luria's approach was useful in localizing functions of the brain as well as finding links between the neural organisation and cognitive organisation. Luria provided syndrome analysis and a qualitative analytic way of finding the underlying neural factors of neuropsychological deficits.

The hypothetico-deductive approach

This approach is based on a theory of the functional organisation of the systems of the brain that regulate neuropsychological functioning (Luria & Majovski, 1977). Qualitative methods have largely been used when studying the relationship between localised lesions and the corresponding psychological and cognitive changes (Luria & Majovski, 1977). This is the clinical paradigm used by UCT neuropsychologists at Groote Schuur Hospital and was central in the design of the current neurocognitive screening battery.

The hypothetico-deductive approach maintains that the neurocognitive function being assessed can be disrupted in a myriad of ways. These different factors can be caused by different determinants, and each factor has complicated structures. Therefore, it is imperative that symptoms must be carefully 'qualified', which depends on a meticulous analysis of the patient's abnormalities (Luria & Majovski, 1977). Luria and Majovski (1977) further elucidate the notion of qualifying symptoms: they understand it as "a clinically creative effort requiring from the neuropsychologist both critical thinking and readiness to reject initial hypotheses if they conflict with new data obtained or if there is confounding of results" (Luria & Majovski, 1977, p.964). Furthermore, the core of qualification is that the clinician needs to observe not only the patient's presence or absence of deficits but rather the kind of dysfunction and the factors that motivate the symptoms. Luria and Majovski (1977) conclude that the salient aspect of assessment is not what aspects are dysfunctional but rather how they are dysfunctional.

Psychometric approach

The psychometric approach, on the other hand, measures cognitive deficits and compares them to test scores that were developed from normative populations. This method aims to show how an individual patient's performance compares statistically to a standardised population (Lezak, 1995). This approach was also partly used in the design of the battery, insofar as patients' test performances are graded with cut-off scores. Predictions are made regarding the nature and location of the lesion based on relationships between tests scores (Lezak, 1995).

The main limitation of the normative approach is that scores are designed in order to make comparisons between groups or within an individual. A score does not properly encompass the test response (Walsh, 1987). Therefore, a wholly quantitative approach would lead to a loss of crucial information; for example, the normative approach would interpret a patient failing a test as the patient having a neuropsychological dysfunction in that cognitive domain. Lezak (1995) goes further and warns that at its worst, a normative approach does not explain the patient's dysfunction. A qualitative approach needs to be integrated in order for meaningful interpretations to be made (Lezak, 1995).

Therefore, a psychometric approach is useful in that it is able to determine whether a cognitive deficit exists. However, it cannot provide the richness of qualitative observations, which provide the information in determining the underlying nature and cause (the meaning) of the abnormality. It is crucial that qualitative judgments be used when interpreting patients' test performances. Walsh (1991) argues that merely

interpreting summarized scores ignores important deficits and that what is needed is a description that conveys “the level of function in the separate areas together with the *nature* of any dysfunctions” (Walsh, 1991, p.258).

What is ideally needed in a neuropsychological assessment is the identification of the precise causes of a deficit and the corresponding brain structures involved with that deficit. More succinctly, psychometric methods are able to determine whether a deficit exists but are not useful in understanding the different systems that are influencing the deficit.

The Groote Schuur Neurocognitive Battery employs a combination of qualitative and psychometric methods with the goal being a theory-driven screening tool. This battery strikes a middle ground between a normative and qualitative approach. The battery uses scores; however, the scores are derived from a hypothetico-deductive line of reasoning (a decision-tree approach). The scores are used to answer yes/no questions regarding whether specific neuropsychological dysfunctions are present. In the Neurocognitive Battery, the Language section is divided into four sections (Production, Comprehension, Repetition and Naming), and the Spatial Cognition section is divided into Spatial Cognition and Perception, Neglect and Anosognosia sub-sections. The focus on each test in each of these sections is to ascertain whether pathology exists in its respective sphere.

Validation in neuropsychological tests

Kaszniak (1989) in Lezak (1995) argues that many neuropsychological tests have the capacity to distinguish between abnormal neuropsychological performances from normal performances. However, these tests have not been developed from large groups and are based on clinical experience (Lezak, 1995). Lezak (1995) provides an example where a valid and reliable and therefore “good” test may be sensitive for visuo-spatial inattention by eliciting no abnormal phenomena from a normal population. However, this same test may only be able to pick up on visuo-spatial abnormalities in some of the patients and will not show reliability when patients’ performances can vary from day to day (Fan et al. 1988 in Lezak 1995). Therefore, statistically sound tests may be of little use for neuropsychological assessment (Lezak, 1995). Graves, Bezeau, Fogarty and Blair (2004) argue that the most common clinical practice is for a test to be able to differentiate between controls and abnormal patients (referring specifically to naming tests).

Mapou (1988) in Lezak further argues that validity will vary in neuropsychological tests depending on where it is applied. Lezak (1995) argues that three types of validity are important. Firstly, the test must be able to measure the neuropsychological abnormalities that the author of the test claims it can (construct validity). It must also display face validity in that colleagues and also patients must have faith that the test appears to measure what it asserts to measure. Predictive validity is referred to as the ability of a test to predict certain criteria based on scores. For example, the scores attained in neuropsychological tests must be able to predict how the patient will perform in natural and authentic situations.

It is important that the sections of the Neurocognitive Battery focused on in this study, namely the Language and Spatial Cognition and Perception sections, show sound validity. The tests as a combination that are part of these respective sections must be able to discern between left and right hemisphere lesions and must discriminate between patients with a neuropsychological dysfunction versus healthy controls.

Neural correlates of spatial cognition and perception

Rapport, Millis and Bonello (1998) maintain that the evaluation of space and object perception is an important component of a neuropsychological battery. De Renzi (1982) in Trojano, et al. (2004) understands visuo-spatial perception as the examination of spatial relationships between objects and with their observer. He argues that this definition is not very specific and limits it further by dividing it into spatial perception and spatial cognition. Spatial perception refers to the elementary processing stages while spatial cognition refers to more intricate mental skills that require internal representations.

Constructional apraxia and visuo-spatial disorders are believed to be related to focal right hemisphere lesions (Trojano et al., 2004). Walsh (1987) agrees and maintains that visuo-spatial lesions can arise from right posterior lesions – the right hemisphere is important in the functioning of spatial perception. Visuo-spatial problems are indicative of a range of neuropsychological disorders, for example, late dementias (Finton, Lucas, Graff-Radford & Uitti, 1998). Heilman (1982) in Walsh (1987) argues that the right hemisphere plays a

large role with regards to attention. Neglect is common in patients with right hemisphere lesions.

Rapport, Millis and Bonello (1998) explain that the Cube Analysis Test can be used for discerning visuo-spatial impairment. They further claim that this impairment is linked with the right hemisphere, with posterior lesions in the right hemisphere leading to the most pervasive impairments. Patients with posterior damage to the right hemisphere often show dysfunction in spatial analysis and object perception tasks. Lezak (1995) argues that the Cube Analysis drawing is a good test for spatial reasoning processes and is performed more poorly on patients with right hemisphere lesions than left hemisphere lesions (Newcombe, 1969; McFie & Zangwill, 1960; Warrington and Rabin, 1970 in Lezak, 1995).

Azouvi et al. (2002) maintain that unilateral neglect is a crucial predictor of weak functional outcome in patients with a right hemisphere stroke. Azouvi et al. (2002) are concerned about the lack of some assessments to assess anosognosia, an associated disorder. Waber and Holmes (1986) in Spreen and Strauss (1991) maintain that the Rey-Osterrieth figure is a helpful test to probe, among other skills, constructional ability.

Neural correlates of language

Diagnosing aphasia is the first crucial step in providing therapy to aid the patient and is commonly assessed through neuropsychological tests. The ICIHD model of the WHO (World Health Organization, 1995 in Kalbe, Reinhold, Brand, Markowitsch and Kessler,

2005) argues that in order to describe language deficit properly, reception and production of verbal and written language, naming and repetition need to be assessed.

Language disorders can be represented in left hemisphere lesions, commonly focal-cortical lesions. Problems with comprehension are usually related to lesions in the posterior left hemisphere and in Wernicke's area. Abnormalities in repetition are associated with lesions bordering the Sylvian fissure. Problems with writing and reading are associated with lesions in the angular gyrus (Devinsky & D'Esposito, 2003). However, it is important to bear in mind that the perisylvian region of the left hemisphere is not the only area responsible for all language processing (Nocentini, Goulet, Roberts & Joannette, 2001).

Global Aphasia

Global aphasia is characterized by poorly articulated speech that is peppered by recurring expressions. Comprehension is severely impaired (Dobel, 1999). All language tests in the Groote Schuur neurocognitive battery give the opportunity to note such behaviour. Notes can be made on repetition, incorrect words used, sentence and grammatical structure. Here, a qualitative approach is preferable in order to make meaningful interpretations of the many functions of the patient.

Broca's and Wernicke's Aphasia

Wernicke's aphasia, or sensory aphasia, is characterized by impaired comprehension (DeAmonet & Thierry, 2001). Neologisms and paraphasias are also very common.

Production of speech is fluent with simple sentence structure (Dobel, 1999). Broca's aphasia, or motor aphasia, however demonstrates agrammatic and non-fluent speech. Modulation is incorrect while comprehension is comparatively preserved (Dobel, 1999). Broca's aphasia comprehension problems are less severe than Wernicke's aphasia, which is especially evident (Bastiaanse & Edwards, 2004). However, Broca's aphasics do not necessarily display normal comprehension (Grodzinsky, Pin˜ango, Zurif, & Drai, 1999).

There is controversy with regard to the precise function of Broca's area, yet it is generally agreed that it is crucial in the production of speech and grammatical functions (Dronkers, 1998). Broca's and Wernicke's aphasia differ in that Broca's aphasia involves grammatical problems and Wernicke's aphasia involves deficits in lexical semantics (Bastiaanse & Edwards, 2004). A characteristic of Broca's aphasia is dysprodic or non-fluent speech where patients exhibit normal levels of speech comprehension but have a tendency to omit function words (Bates, Friederici & Wulfreck, 1987).

Boston Naming Test

The most popular confrontation naming test is the Boston Naming Test, which was developed in 1983 by Kaplan, Goodglass and Weintraub (Marien, Mampaey, Vervaeet & Saerens, 1998). It is the most frequently used naming test in the West, especially North America (Barker-Collo, 2001; Kohnert, Hernansez & Bates, 1998; Spreen & Strauss, 1991; Welch, Doineau, Johnson & King, 1996). The test is comprised of sixty ink sketches which vary in complexity from a house to an abacus. Patients are required to name each object depicted in the drawing. This is a crucial test for aphasics, as naming is

generally agreed as a delicate facet of underlying aphasic dysfunctions (Welch, Doineau, Johnson & King, 1996), and anomia or word-finding difficulty is the most common sign of aphasia (Ardila, 2007).

The Boston Naming Test's functions are impressive as it can discern naming problems from aphasics (Margolin, Pare, et al., 1990 in Lezak, 1995) and is successful in detecting the varying levels of dementia (Lezak, 1995) and multiple sclerosis (Lezak, Whitman and Bourdette, 1990 in Lezak, 1995). It is also helpful in evaluating other injuries and diseases, from progressive neurodegenerative diseases, brain injuries, and cerebrovascular accidents (CVAs) to anoxia (Kim & Na, 1999; Marien, Mampaey, Vervaeke & Saerens, 1998).

Concerns with Boston Naming Test

Ardila (2007) mentions that there are several limitations with the BNT. He argues that because of its development in an American context, it is biased. For example, many of the items are American-specific elements (e.g. beaver). Ardila (2007) warns that one's naming ability is strongly related to one's level of education. Kim and Na (1999) acknowledge that there are no Korean norms for many neuropsychological tests and therefore were urged to develop their own version of the Boston Naming Test specific to their country. Many of the items in the BNT were not suitable for a non-Western population. They reported problems with items such as "beaver", "pelican", etc. that were similar to the many problem items that our pilot study raised.

Barker-Collo (2001) argues that the BNT is sensitive to cultural experience. Yiu, Hickson, and Barnett (1995) in Barker-Collo (2001) claim that culturally inappropriate items in the BNT need to be replaced. Barker-Collo (2001) suggests that the BNT may not be applicable outside of the culture for which it was intended and that a modified version might be helpful. They also suggest that one version of the BNT cannot be generalised for an entire country but only more for a specific subset of the population. Worrall, Yiu, Hickson and Barnett (1995) in Kim and Na (1999) argue that the norms developed for mainly North American populations cannot be used in other areas.

Fastenau, Denburg and Mauer (1998) state that short forms of neuropsychological tests such as the Boston Naming Test provide many advantages. They argue that in a clinical setting, long batteries result in excessive frustration in the patients. Graves, Bezeau, Fogarty and Blair (2004) suggest that a shortened form of the BNT would be helpful, especially for screening purposes and assessing patients with limited concentration and motivation, and it saves time. AERA, et al. (1999) in Graves, Bezeau, Fogarty and Blair (2004) assert that it is crucial to establish test validity with a specific use of the test in mind.

Ethnicity also influences naming performance, yet it has not been thoroughly investigated (Randolph, Lansing, Ivnik, Cullum & Hermann, 1999). There is a large variability in the performances of people from different cultures and races (Hamsher, 1984 in Randolph et al., 1999). For example, in New Zealand, the Maori population scores significantly more poorly than Europeans (Barker-Collo, 2001).

At Groote Schuur Hospital, the neuropsychologists have identified consistent problems when using the Boston Naming Test with South Africans. They found items such as 'pretzel', 'trellis' and 'beaver' to be particularly unsuitable for South African patients. The problem arises when the neuropsychologist is unable to discern whether an incorrect answer is due to culturally biased items or cognitive deficit. For example, it becomes hard to judge whether the patient is suffering from semantic paraphasias or is healthy but strained to answer culturally biased questions.

3.) Aim

The aim of the present study is, firstly, to rectify remaining problems with newly adapted tests. Certain items of the Naming Test and the Cube Analysis Test need to be re-adapted and piloted on a group of control participants. In addition, a scoring sheet that will accompany the assessments using the full Neurocognitive Battery needs to be developed before the validation process can begin in order to note performances and record qualitative observations. The final step is to validate the neurocognitive battery: the battery needs to be able to discriminate between abnormal and normal neuropsychological performances.

The method taken in validating the battery is that it needs to demonstrate that patients' test performances relate accurately to known anatomical lesions and elicit test performances that are characteristic of those specific lesions. Each patient with left- or right- hemisphere lesions was given the whole battery. The battery must show that those with left-hemisphere lesions perform more poorly on the language tests than on the

visuo-spatial aspect of the battery and those with right-hemisphere lesions perform worse on the visuo-spatial tests than the language tests.

The battery must also show reliability. The battery must show adequate test-retest reliability and inter-rater reliability – in other words, the battery needs to be consistent and to attempt objectivity (Durrheim, 1999). A battery cannot show validity if it is not demonstrated to be consistent over time (test-retest reliability). The battery should be consistent irrespective of the trained neuropsychologist using it (inter-rater reliability).

In summary, the goal of this research is to assist with the development of a theory-driven, neurocognitive screening battery (which has equivalent forms in English, Afrikaans and isiXhosa) by conducting a study which contributes meaningfully to the overall improvement of the tests and helps with the validation process.

4.) Method:

Re-piloting Phase

Sample

The sample for the second phase of piloting of the neurocognitive tests consisted of thirty neurologically healthy, cognitively intact control participants. The majority of the sample was enlisted from Groote Schuur Hospital in order to be representative of the patient population seen clinically. All participants were screened beforehand to exclude possible neurocognitive disorders; anyone with a history of neurological and psychiatric disorders, metabolic dysfunctions, head injury and/or other diseases was excluded (see Appendix One).

This sample of thirty was comprised of an equal number of Afrikaans, isiXhosa and English speakers (i.e. ten participants for each language group). The average age of the participants was forty-two years old, with a minimum age of sixteen and a maximum age of 74. The sample comprised of eighteen males and twelve females. See Table 1 for further demographic information.

Table 1.

Means and standard deviations for Age and Years of Education for isiXhosa, English and Afrikaans groups

Controls	Language		
	IsiXhosa	English	Afrikaans
	Age		
<i>M</i>	35.6	49.7	44.22
<i>SD</i>	12.31	12.75	17.21
	Years of Education		
<i>M</i>	8.4	10.2	7.8
<i>SD</i>	2.91	1.93	3.35

Materials

The materials used included the original versions of the Naming Test and the 3-D Analysis Test (described below) created from the first phase of adapting, which required further adaptation, as well as these two tests in their latest form, following this present study's further item modifications ahead of the re-piloting phase. A scoring sheet

specially developed to score responses and to make qualitative notes on the context and nature of the participant's responses was also used (see Appendix Four).

Other materials used for this phase of the study were a Consent Form and Participant Information Sheet, created to gain permission to test the participants and explain the purposes of the study (see Appendices Two and Three); and finally, the Screening Sheet was designed to exclude those with possible pre-morbid neurocognitive disorders and thereby ensure a healthy control sample (see Appendix One).

Design

The design of this study was geared towards establishing the effectiveness of the changes made to the newly developed neurocognitive tests by re-piloting the tests that still required further adaptations and then drawing on converging lines of evidence, both qualitative and quantitative data, to evaluate the efficacy of the changes. As mentioned in the introduction above, a combination of both qualitative and quantitative data is argued to be the most successful approach to neuropsychological tests (see Akhuntingina & Tsvetkova, 1983). Quantitative analysis was beneficial to determine whether or not deficits existed and qualitative analysis was able to provide the nature of the underlying causes of the deficit.

The converging lines of evidence took the form of feedback from neuropsychologists; the language and cultural experts previously used on other aspects of the overarching research; and finally from the participants themselves, their test performances and their

qualitative feedback as to which tests/items they struggled with and why they did so. Qualitative data was gathered from the patients' preferences, from their new versus original test performances, and from the feedback from the neuropsychology clinicians and the language and cultural experts. Quantitative data was gathered from the participants' test performances.

From this multifaceted process, it could then be established which tests and test items were suitable for final inclusion into the Neurocognitive Battery prior to the commencing of the validation phase of the study. This initial re-piloting phase of the study was not a single-blind design, as it was known that all the participants were controls.

Data Analysis

Data analysis was approached from both a qualitative and quantitative perspective. The four new Naming Test items needed to be evaluated to determine whether participants were familiar with the items and that the three language groups performed similarly. Chi-squared tests and t-tests were used to make comparisons between the three language groups.

Factorial ANOVA was used to analyse the efficacy of the new Naming Test after the two problematic items were replaced by the two most suitable from the new items. This was done by determining which variables, such as Age, Language and Level of Education, influence controls' performance on the Naming Test.

For the 3-D Analysis Test, t-tests were used to determine the difference in performance between the old and new version of the 3-D Analysis Test. Factorial ANOVA was used to determine what variables, such as Age, Language and Level of Education, influence test performances.

Data Collection Procedure

Further adaptation of the Naming Test. The previous phase of the overall study had highlighted many problematic items with the original Boston Naming Test (BNT). The odd-numbered, short version comprising of thirty BNT items was used. As a result, all thirty BNT test items were replaced and alternate items were constructed, with additional input and guidance from the language and cultural experts at the University of Cape Town as well as the neuropsychologists at Groote Schuur Hospital. The cultural experts were needed in order to discern whether items could be identified by all three targeted South African language groups (English, Afrikaans and isiXhosa) and to ascertain their level of difficulty in each language group, especially in populations of low socio-economic status.

The neuropsychologists' input was needed to confirm whether items were complex enough to be able to detect neurocognitive dysfunctions such as semantic paraphasias. A paraphasia refers to the incorrect replacement of words or syllables. A semantic paraphasia is the replacement of a word with one that is semantically related, such as "John drove the wheel" (Devinsky & D'Esposito, 2003: 174), where "wheel" replaces "car". These are very common in language disorders or aphasias. The goal was to find

South African equivalents to the Boston Naming Test items; for example, the item, 'ostrich' replaced 'pelican'.

After the first phase of the study replaced the BNT with a new thirty-item test (the Naming Test) and piloted them on 30 controls and three aphasic patients, a few problematic items were identified with this new Naming Test. These items were the 'hippopotamus', the 'mug', and the 'dragonfly'. Therefore, replacements were needed for these problematic items. The items were problematic because most of the isiXhosa-speaking controls could not name these items correctly, as their colloquial language does not discriminate between 'cup' and 'mug' and they frequently mistook the 'hippopotamus' for a 'pig' as hippopotami are not found in the Eastern Cape where many isiXhosa people live before moving to Cape Town.

The items 'mug' and 'hippopotamus' were replaced by the items 'pylon' and 'aloe'. The rationale behind these items was that the Naming Test also required one or two more difficult items that not everyone would be able to name but that would be familiar to all three cultures and those living in urban or rural areas. This recommendation for the addition of complexity was based on another minor problem identified from the initial pilot study on the Naming Test, as the results showed that the overall average score for the new test presented a slight ceiling effect relative to international average scores. It was decided to keep the 'dragonfly' item because while many participants were unable to name it correctly, they did report recognizing it and being familiar with it, and additionally, one or two more difficult items were required.

The 'aloe' plant is commonly known in the Western Cape and 'pylons' are found all over the country, yet while these items are familiar, people often have difficulty naming them correctly, according to the language and cultural experts at the University of Cape Town. Furthermore, all three languages have terms for these items, and the cultural experts concluded that these items were difficult to name but were known by urban and rural people and Afrikaans, isiXhosa and English populations. Other items were also included as possible additions to the Naming Test, resulting in the new items added to the Naming Test being the following: 'pylon', 'aloe', 'spur' and 'husk' (of corn).

Further adaptation of the 3-D Analysis Test. It was found that controls in the previous project had performed quite poorly on the newly created 3-D Analysis Test and its predecessor, the Cube Analysis Test. The 3-D Analysis Test involves showing participants 2-D representations of various stacked cube configurations and asking them to count all the cubes, including hidden ones, that make up the configuration. Problems with the test ranged from the blocks being too small for the elder participants with eyesight problems, as well as the cultural and education bias of visuo-spatial tests (Rossellia & Ardila, 2003). Additionally, from the qualitative feedback from the initial pilot study, many of the participants had complained that the blocks in the test were distorted in perspective (Mosdell, 2005), a critical observation which the neuropsychology experts had also made. The cubes were distorted in that the perspective of the cubes was not drawn correctly, leading to a confusing representation.

It was necessary to simplify the 3-D block design and to enlarge the size of the blocks. Attempts were made by using artist drawings of the block designs. Eventually, the most appropriate result was made through a computer programme called Blender, which is used as a 3-D computer generated image application. Simplifying the 3-D block design was completed with the help of neuropsychologists at Groote Schuur hospital, and fourteen block designs were then finalised.

Items were identified that the majority of the controls, from the previous study, answered correctly. These items were retained. Items that were performed poorly by controls were isolated and replaced with less complicated block designs with fewer blocks. All block designs were created and enlarged with the aforementioned Blender application. Finally, the neuropsychologists from Groote Schuur Hospital were consulted, and their feedback was needed to ensure that the blocks in the 3-D Analysis Test were simplified enough before testing them on controls, yet would still be sensitive enough in identifying space perception difficulties (personal communication with Prof. M. Solms).

Administration of tests. All thirty participants granted their consent. The newly modified Naming Test and 3-D Analysis Test were administered to the controls. The instructions from the battery were given to the participants verbally. Instructions were administered in the first language of each individual participant as a previous phase of the overall project had translated and back-translated the battery and its tests into isiXhosa and Afrikaans from its original English.

Both the original and the new test versions were administered together in one session per participant. The average length of the testing session was 20 minutes because tests from other sections of the Groote Schuur Battery (memory and executive sections) were administered during the session as part of the larger single-blind design, discussed below. After each session, participants were thanked for their contribution to the study and their time. Those participants who had to travel to the hospital received R100 gratuity for their time.

The specially developed scoring sheet was used to record participants' responses. The examiners for the pilot study were masters' students from the Neuropsychology Division at Groote Schuur Hospital and were researchers for this study who received prior training in the scoring procedures of the battery and test administration. These examiners had also been involved in the previous piloting phase as part of the larger overarching study.

In addition, an isiXhosa or Afrikaans interpreter accompanied the researchers when assessing participants who fell within these respective language groups. These interpreters had prior experience working with patients at Groote Schuur Hospital had served as the translators for the original Groote Schuur Neurocognitive Battery, and were members of the panel of cultural and language experts assembled for the larger study. They had also been involved in the initial pilot study and had thus already been trained by a neuropsychologist in the administration of the Neurocognitive Battery and its tests.

Validation Phase

Sample

Sampling consisted of selecting suitable patients with established lesions resulting in neurocognitive deficit, who were admitted to Groote Schuur Hospital. This is a form of purposive or judgmental sampling (van Vuuren & Maree, 1999), in that the neuropsychologists selected the cases with specific characteristics in mind. Patients were selected on two grounds, firstly on the basis of lesion site, and secondly on the basis that they presented with typical neurocognitive deficits.

It is impossible to be able to conduct a thoroughly randomised sampling for this study, as it is necessary to find patients with specific neuropsychological dysfunctions. Ideally, one would adopt a randomised sampling technique in order to ensure an unbiased representation of one's population. However, the type of population that this battery targets are the patients at Groote Schuur with potential neurocognitive deficits. These patients are difficult to find in large numbers, and therefore any patients who fit the criteria for this study at Groote Schuur Hospital were approached. These patients' symptoms can also quickly resolve, which further limits the available population.

Criteria for patients in the left-sided stroke group needed to present with Left Middle Cerebral Artery ((L) MCA) or Left Internal Capsule cortical extension strokes and associated language dysfunctions. Right-sided stroke patients needed to present with Right Middle Cerebral Artery ((R) MCA) strokes with corresponding neurocognitive symptoms involving visuo-spatial impairment.

In total, ninety participants were assessed between two researchers. These ninety participants comprised equal numbers of left-sided stroke patients (aphasics), right-sided stroke patients, amnesiacs, patients with executive dysfunction, and control participants (fifteen participants in each group). However, the focus of this study is on the forty-five left-sided stroke patients, right-sided stroke patients and control participants (fifteen participants from each group). The participants ranged in age from nineteen to eighty-four years old with a mean of forty-two years old. The participants' levels of education ranged from zero to sixteen years with a mean of approximately nine years. The sample consisted of eighteen males and twelve females. See Table 2 for further demographic information.

Table 2.

Means and standard deviations for Age and Years of Education for isiXhosa, English and Afrikaans groups

Validation Group	Language		
	IsiXhosa	English	Afrikaans
	Age		
<i>M</i>	49.7	44.22	42.23
<i>SD</i>	12.75	17.21	14.97
	Years of Education		
<i>M</i>	10.2	7.8	8.8
<i>SD</i>	1.93	3.35	2.89

Materials

The new Neurocognitive Battery, including the Naming Test and 3D Analysis Test, was administered in its entirety with all three language versions available, namely, IsiXhosa, Afrikaans and English. The Language section of the battery consists of a Naming section, which includes the Naming Test to assess confrontational naming, questions asking the patient to name body parts and items that are found around the patient's bed; a Repetition section, which includes sentences varying in complexity to ascertain repetition ability; a

Comprehension section with questions that assess comprehension ability which range from a simple command (“close your eyes”) to asking participants what their relation is to family members (“what is your relation to your sister’s daughter”); and the Token Test (Boller & Vignolo, 1966; De Renzi & Vignolo, 1962 in Lezak, 1995), which is an option test in the Neurocognitive Battery and does not form part of the main scoring protocol.

The Naming Test is also an optional, more complex test and is scored separately to the main battery. The Language section also includes a Production section that includes a passage for the participant to read (the Mary Selo story); a writing section where the patient is asked to write his or her name, write a sentence from dictation and write a spontaneous sentence, and a narrative production test, the “Washing Line Picture Test”.

The Washing Line Picture Test was based on the Cookie Theft picture from the Boston Diagnostic Aphasia Examination (Goodglass & Kaplan, 1983 in Lezak, 1995). However, the Washing Line Picture Test was developed to reduce cultural bias and was piloted on South African controls and aphasics two years ago (Mosdell, 2005). The Mary Selo Story, a reading test that contains South African-specific locations and names, was developed at the same time. See Appendix Four for the complete Language section of the Neurocognitive Battery.

The Spatial Cognition section of the battery comprises of Spatial Cognition and Perception, Neglect and Anosognosia sections. The 3-D Analysis Test, the Rey-Osterreith Complex Figure (Osterreith, 1944 in Lezak, 1995) and three arithmetic

questions form the Spatial Cognition section. For the Anosognosia section, questions are present that ask patients about their symptoms/deficits. The Neglect section is comprised of questions to assess unilateral neglect in three modalities (auditory, tactile and visual) and the “Hut Drawing Test”, which is a drawing for the participant to copy. The Hut Drawing Test was developed and piloted two years ago, along with the Mary Selo story, Washing Line Picture Test and first version of the Naming Test (see Appendix Five and Six for the complete Spatial Cognition and Language section).

The scoring sheet, mentioned previously, which was specially developed to make qualitative notes on each patient and record scores was also used. The Consent Form and the Patient Information Sheet were used to explain the purposes of the study to patients and to gain their permission (see Appendices Two and Three). These documents were the same ones used in the re-piloting phase.

Research design

A single-blind design was used for the validation phase. Although the neuropsychologist was aware of whether each participant was a patient, and of which type, or a control, the researchers collecting the data were unaware of whether their participants were left or right-sided stroke patients, amnesiacs, executive dysfunction patients or controls. All control participants used in this study were screened for any neurocognitive deficits by a neuropsychology staff member. It is crucial to ensure that participants with any neurological problems are not used as control participants; the control sample used must

reflect neurologically healthy individuals in order to be able to test accurately that the battery does not suggest neurocognitive deficits in the absence of pathology.

Three participants from each of the five participant groups (left-sided stroke, right-sided stroke, executive dysfunction, amnesiacs and controls) were randomly selected to be re-assessed by another researcher who had not performed the initial assessment in order to determine the test-retest and inter-rater reliability of the battery. To reiterate, the focus of this study was to assess the validity and reliability of the Language and Spatial Cognition sections of the battery, i.e. to test whether the battery can accurately discern between left and right hemisphere lesions, and whether it can do so consistently.

Once the data collection process was completed, the data that each researcher was responsible for would be handed to him or her (either language and visuo-spatial participants or executive and memory participants), and whilst still blind as to the participants' actual groups, the researcher would allocate the participants to one of three possible groups his or her study was responsible for, i.e. either one of the two pathological groups, in this case, the language or visuo-spatial dysfunction groups, or the control group.

This allocation would be done solely on the basis of the qualitative observations of the participants' performances on each specific test rather than by looking at their overall battery score or its sub-scores. This battery's development and scoring procedures were developed from a qualitative and quantitative perspective, and it is necessary to evaluate

it from both perspectives. This process was repeated by another researcher so that inter-rater reliability could be assessed. Once this qualitative process was over, the researchers were 'un-blinded' and the statistical procedures (see below) could be conducted for the quantitative aspect of the validation study. This design ensures that the validity of the battery can be meaningfully ascertained using both qualitative and quantitative findings.

Data Analysis

As previously mentioned, the data analysis comprised of both qualitative and quantitative techniques. This combination of approaches was necessary to fully evaluate the utility of the tests in the screening battery without losing any crucial information regarding the participants' performances.

Qualitative analysis. After the assessments were complete, both examiners, while still blind to the participants' pathology/control status, received all forty-five of the scoring sheets. For this study, these forty-five scoring sheets comprised of controls, left hemisphere lesions or right hemisphere lesions assessments. The examiners were then required to categorise their forty-five scoring sheets into three groups, using the patients' performances on the tests and any qualitative comments to judge to which group, control, language or visuo-spatial dysfunctions, each scoring sheet belonged. Once the two researchers had each allocated his or her own, and then the other examiner's data into the three respective groups, the discrepancy between how they allocated the participants and the actual groups the participants were from was analysed. The problems that the scorers encountered were investigated, as well as factors that aided diagnosis.

Quantitative analysis. Quantitative analysis focused on one-way ANOVA in order to ascertain whether the battery could discern between the left hemisphere lesion, right hemisphere lesion and control groups. Factorial ANOVA was used to determine the factors, such as Age, Language and Education Level, that influence the performance on the tests in the battery. The total scores of the Language and Spatial Cognition section as well as individual test scores for those two domains were examined by using Factorial ANOVA to determine each test's ability to discern pathology and recognise healthy participants. Chi-squared contingency tables were also made for the categorical data to determine which factors were associated with performances on Spatial Cognition section tests.

Procedure

The entire Neurocognitive Battery, including the Language and Spatial Cognition sections, was administered to controls and patients, with the examiner and interpreter (when required) being blind as to which group each participant was from. Standardised verbal instructions of the battery were given to all participants. The specially developed scoring sheet that was used in the re-piloting section of the study was used here to record participants' responses.

Participants were assessed in their first language. Therefore, when necessary, an interpreter, either Afrikaans or isiXhosa, would accompany the researcher. Researchers were the same masters' students who had tested the controls in the re-piloting section of the study. They therefore had been trained in test administration and scoring procedures.

Interpreters used in the re-piloting section of the study were the same ones used in this validation phase and were trained in the test administration of the battery and were also involved in translating the battery into its Afrikaans and isiXhosa forms.

The entire battery was administered in one testing session to each participant. The testing session lasted approximately sixty to ninety minutes, as the entire Neurocognitive battery was administered in a single session. Assessments took place at Groote Schuur Hospital. Three participants from each of the five groups (left-sided stroke, right-sided stroke, amnesiacs, executive dysfunctions and control groups) were re-assessed by another researcher for reliability purposes. These re-assessments were carried out within two days of the first assessment, as neuropsychological symptoms can resolve quickly.

Ethical considerations

Informed consent was obtained from all participants (see Appendix Four). Participants' requirements were made unambiguously clear. The requirements of the study needed to be explained in a non-technical way so that any potential participant could understand. Interpreters were used in order to ensure that all participants received information in their first language. Durrheim and Wassenaar (1999) argue that information must be available to the participant even after the assessment has begun, and the participants were reminded that they could ask for clarification at any time.

Patients were informed that they could discontinue the study at any time without any negative consequences incurred upon them. This was important since they were currently

in a hospital being treated and may have needed reassurance that they would still receive proper medical care even if they declined to take part or wanted to discontinue participation during the assessment.

Responses were confidential; they were only made available to the research team and the interpreters. Explanations were made to the participants that their responses would be safely stored and that their personal identity would remain concealed to ensure their privacy. This was paramount because medical information was gathered from them if they agreed to take part. Only pertinent information for the study was gathered from the participants. Ethical permission for this study was granted by the Groote Schuur Research Ethics Committee and by the University of Cape Town Psychology Department's Ethics Committee.

5.) Results

Re-piloting phase

Naming Test

Participants performed similarly on the four added items to the Naming Test, with twenty-two incorrect responses for 'pylon', twenty-one incorrect scores for 'spur', fifteen incorrect scores for 'husk' and fourteen incorrect scores for 'aloe'. Chi-squared tests were performed to determine whether the four items showed language bias. The number of correct/incorrect responses did not differ by language group for each of the four test items - for the Pylon $\chi^2(2, N = 30) = 4.42, p > .05$, Spur $\chi^2(2, N = 30) = 2.85$, Husk $\chi^2(2, N = 30) = 2.4$ and Aloe $\chi^2(2, N = 30) = 1.07$ items. A significance level of $\alpha = 0.05$ was used for all analyses.

Qualitatively, participants who did not name the objects correctly maintained that, while they were familiar with the 'pylon' and 'aloe' items, even though they had difficulty recalling them, the 'husk' and 'spur' items were unfamiliar. Some participants were not aware that there were specific terms for these items in their first language. Common responses to the latter items were 'leaves' or '*blare*' (Afrikaans) for 'husk' and 'claw' for the 'spur'.

Further analyses were made after the problematic 'mug' and 'hippopotamus' items were replaced by the 'pylon' and 'aloe' items, resulting in an adjusted thirty-item Naming Test. The average score for the new Naming Test was 25.97 with a standard deviation of 3.13 (see Table 3 for the language breakdown of descriptive statistics).

Table 3

Means and standard deviations of Naming Test scores for isiXhosa, English and Afrikaans groups

	Language			
	IsiXhosa	English	Afrikaans	Total
	Naming Test			
<i>M</i>	25.9	26.2	25.8	25.97
<i>SD</i>	3.21	3.46	3.05	3.13

A cell mean plot was created in order to show how the language groups performed on the new Naming Test. Below, Figure 1 shows that the three language groups performed very similarly to one another. One-way ANOVA was calculated to determine the actual probability that the three languages' performances resulted from the same population. There was no significant effect, $F(2, 27) = 0.04, p = 0.96$. This is an extremely high probability. Therefore, the three language groups perform very alike on the new Naming Test.

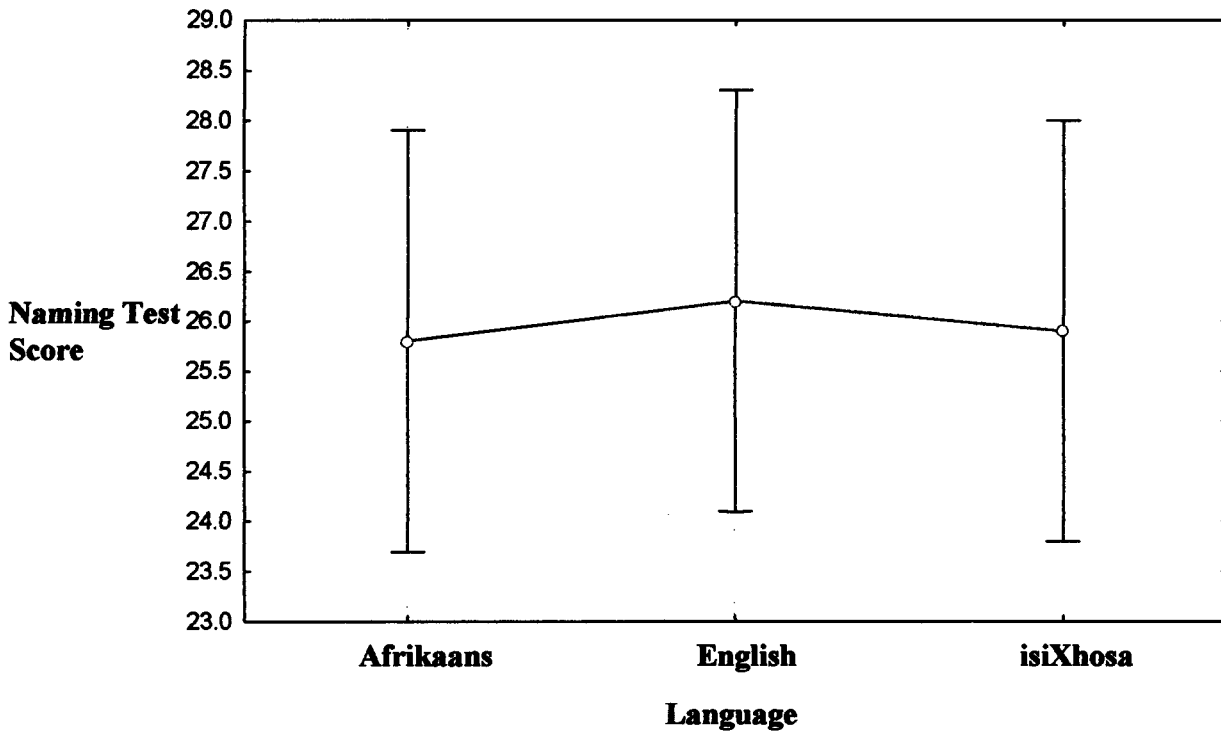


Figure 1. Cell mean plot showing Naming Test performances for each language group

Factorial ANOVA was performed to determine the influence level of education on the new Naming Test performance and how it interacted with language in the sample. Assumptions of normality and homogeneity of variance were upheld. Table 4 shows the descriptive statistics (means and standard deviations) of the analysis. There was no significant main effect for language, $F(1, 39) = 0.62, p = 0.54566$, but there was a significant effect for education, $F(2, 39) = 11.57, p = 0.002349$. There was no significant interaction effect, $F(2, 39) = 0.43, p = 0.654841$.

Due to the relatively small size of the sample, education level was broken into two groups: those with less or equal to eight years of education and those with above eight years of education. This is a logical division, as this is the cut-off between primary/junior school and high school. This also allows the data to be divided as equally as possible, as the mean education level of the sample is 8.8 years, which ensures equal sample size.

Table 4. Mean naming test score.

Factor A: Education	Factor B: Language			A marginal means
	English	Afrikaans	isiXhosa	
8 years or more	¹ 27.25	28.00	27.20	27.41
	² 2.55	0.82	2.05	
	³ 8	4	5	
less than 8 years	¹ 22.00	24.33	24.60	24.08
	² 4.24	3.14	3.85	
	³ 2	6	5	
B marginal means	26.20	25.80	25.90	

¹ Mean ² St. Dev ³ N

A t-test was calculated to ascertain the influence that age had on the new Naming Test's performances. Thirteen of the participants were forty years old or younger, and seventeen of the participants were older than forty years. There was a significant effect for age, $t(28) = 2.64, p = 0.013508$, with the younger group ($M = 27.54$) scoring higher than the

older group ($M = 24.76$). This shows that participants older than forty years old perform worse on the new Naming Test than participants forty years old and younger.

3-D Analysis test

Comparisons were made between the participants' scores on the adapted Cube Analysis Test and the new 3-D Analysis Test. A t-test was used to determine whether the two versions of the test performed differently. There was no significant difference between new 3-D Analysis Test scores ($M = 8.63$) and old 3-D Analysis Test scores ($M = 8.1$), $t(28) = 0.49$, $p = 0.619988$. This shows that there was not a significant difference between how control participants performed on the two tests. See Table 5 for the new 3-D Analysis Test's descriptive statistics by language group.

Table 5.

Means and standard deviations of 14 item 3-D Analysis Test scores for isiXhosa, English and Afrikaans groups

	Language			
	IsiXhosa	English	Afrikaans	Total
	3-D Analysis Test			
<i>M</i>	8.7	10.8	6.4	8.63
<i>SD</i>	3.02	4.34	4.27	4.21

Factorial ANOVA was used in order to determine whether language and education level influenced performance on the new version of the 3-D Analysis Test. Assumptions of normality and homogeneity of variance were upheld. Education level was divided into two groups: those with eight and below years of education and those with above eight years education. In Table 6, descriptive statistics (means and standard deviations) are shown for the analysis. There was a significant main effect for education, $F(1, 39) = 16.28, p = 0.000482$ but no significant effect for language, $F(2, 39) = 0.88, p = 0.426888$. There was a significant interaction effect, $F(2, 39) = 3.44, p = 0.04845$. This shows that Language does not influence 3-D Analysis Test performance. Therefore, all three language groups performed similarly.

Table 6. Mean score for new 3-D Analysis Test

Factor A: Education	Factor B: Language			A marginal means
	English	Afrikaans	isiXhosa	
8 years or more	¹ 12.63	9.25	9.2	10.82
	² 1.30	3.59	3.11	
	³ 8	4	5	
less than 8 years	¹ 3.5	4.5	8.2	5.77
	² 4.95	3.78	3.19	
	³ 2	6	5	
B marginal means	10.8	6.4	8.7	

¹ Mean ² St. Dev ³ N

However, Level of Education does impact performance on the new 3-D Analysis Test. Cell mean plots (see Figure 2) show that those participants with eight years of education or below perform significantly less well on the new 3-D Analysis Test than those with more than eight years of education. The Factorial ANOVA analysis also showed that there was a significant interaction between Level of Education and Language.

Tukey's HSD (Honestly significant difference) showed where the significant difference lay: between the English group with a high education level and the Afrikaans group with a low education level ($p = 0.0008$). There was also a significant difference within the English group between the high and low education levels ($p = 0.011085$). The cell mean plot in Figure 2 shows that the interaction effect occurs at the isiXhosa language level. IsiXhosa-speaking participants were not influenced by education level to the degree of their Afrikaans and English-speaking counterparts. While the cell mean plot clearly shows that the Afrikaans and English-speaking groups perform much lower on the new 3-D Analysis Test if they fall into the lower education group, there is no such disparity in the isiXhosa-speaking group.

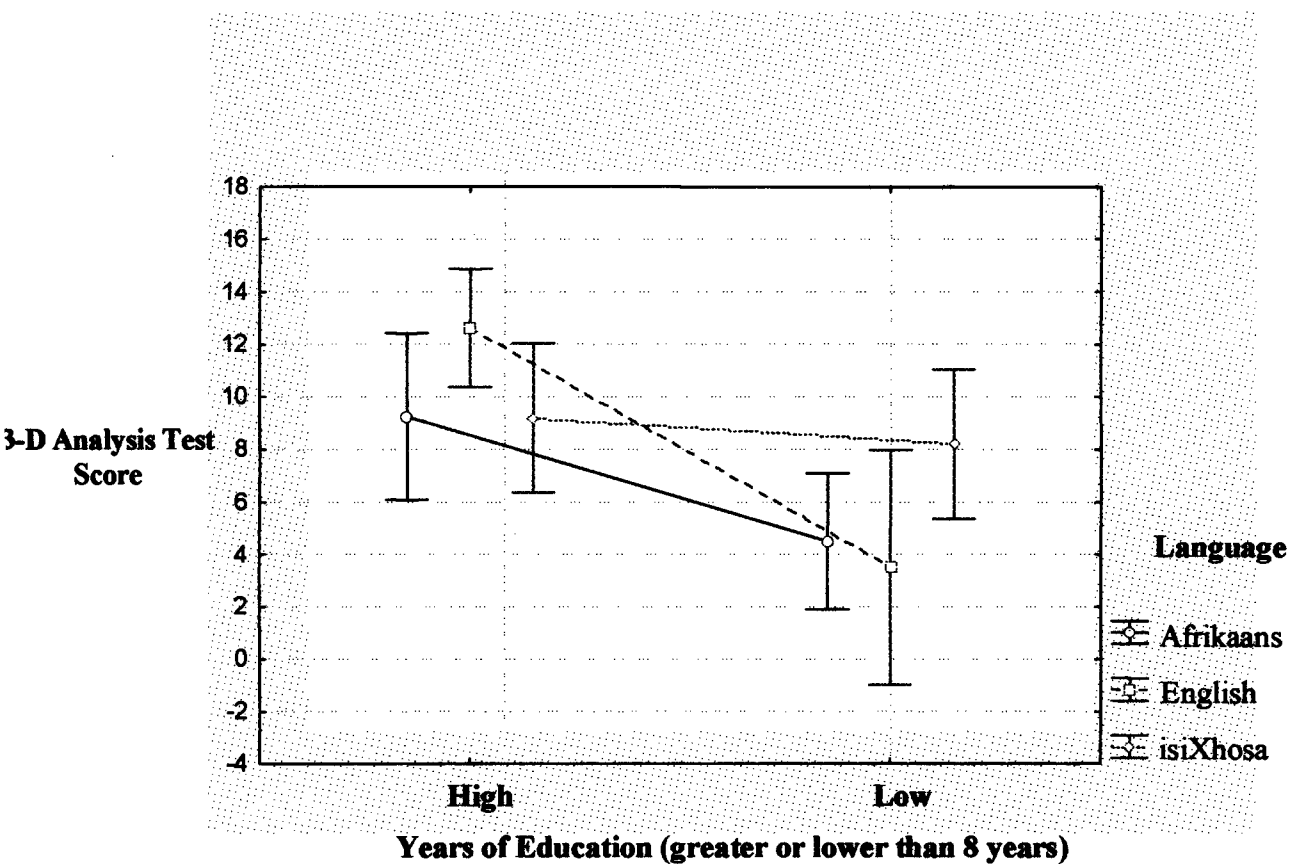


Figure 2. Cell mean plot showing 3-D Analysis test performances by language group and education level

An independent sample t-test was performed to ascertain whether age also influenced new 3-D Analysis Test performance. Participants were divided into two groups: those forty-five years and older and those who were younger than forty-five years old. The forty-five year old cut-off divides the sample equally (sixteen in the younger group and fourteen in the older group), which would allow future ANOVA tests to be performed if the assumption of homogeneity of variance is violated, by keeping sample size as equal

as possible. There was no significant effect for age, $t(28) = 0.85, p = 0.4$. The probability value was insignificant, which shows that age did not impact on the control participants' performances.

Calculations were made to determine which of the fourteen items from the new 3-D Analysis Tests were most frequently incorrectly answered. More than a third of the control participants failed six out of the total fourteen items in the test. From these six poorly answered items, four were removed in order to make the new 3-D Analysis Test easier.

With the four items removed from the 3-D Analysis Test, the average performance increased to 7.03 (out of 10) from 8.63 (out of 14) which would be equivalent to 6.16 out of 10 (see Table 7 for descriptive statistics). A factorial ANOVA was performed to determine the influence of Language and Education on the new 10-item 3-D Analysis Test. See Table 8 for descriptive statistics (means and standard deviations). There was a significant main effect for education, $F(1, 39) = 16.72, p = 0.000420$, but no significant effect for language, $F(2, 39) = 1.29, p = 0.293513$. There was no significant interaction effect, $F(2, 39) = 2.21, p = 0.131052$.

This shows that language did not have an influence on participants' performances, and there was no interaction effect between level of education and language. Level of education, however, was found to be highly significant. The participants' level of education influences their performance on the 10-item 3-D Analysis Test.

Table 7.

Means and standard deviations of 10 item 3-D Analysis Test scores for isiXhosa, English and Afrikaans groups

	Language			
	IsiXhosa	English	Afrikaans	Total
	3-D Analysis Test			
<i>M</i>	7.5	8.1	5.5	7.03
<i>SD</i>	2.64	3.11	3.1	3.07

Table 8. Mean score for 10-item version of new 3-D Analysis Test

Factor A: Education		Factor B: Language			A marginal means
		English	Afrikaans	isiXhosa	
8 years or more	¹	9.38	7.75	8.20	8.65
	²	0.74	1.71	2.49	
	³	8	4	5	
less than 8 years	¹	3.00	4.00	6.80	4.92
	²	4.24	2.97	2.86	
	³	2	6	5	
B marginal means		8.10	5.50	7.50	
¹ Mean	² St. Dev	³ N			

Validation and reliability phase

Reliability

Inter-rater and test-retest reliability of the Language and Spatial Cognition sections of the Neurocognitive Screening Battery was determined by re-assessing three participants from each of the three participant groups (left hemisphere lesion, right hemisphere lesion and control). Re-assessments were completed by a different researcher who was also blind as to which of the three groups each participant was from. Therefore, in total, nine re-assessments were done. Four of the participants were English-speaking, two were isiXhosa-speaking and two were Afrikaans-speaking. The battery was administered to each of them twice within a maximum of a forty-eight-hour interval in order to minimise the impact of neurological patients' resolving on test-rest reliability (see Table 9 for further demographic characteristics of re-tested participants).

Table 9.

Means and standard deviations of Age and Years of Education for reliability participants. (N=9, Male=2, Female=7)

Age	
<i>M</i>	49.67
<i>SD</i>	11.48

Years of Education	
<i>M</i>	8.5
<i>SD</i>	4.98

Pearson's product moment correlation coefficients (*r*) were made between individual scores on the Orientation section of the battery, the Language section and the Spatial Cognition section. A further correlation was made of the summation of these sections.

The participants' individual scores on the tests were first converted to percentages in order to normalise the scores. The summations of these individual scores were then correlated with the re-test assessments. See Table 10 for correlation coefficients.

Table 10

Reliability correlation coefficient for Orientation, Language and Right-Hemisphere sections

Section	Correlation Coefficient
<i>Total section (Orientation, Language and Right-Hemisphere section)</i>	0.89
<i>Orientation section</i>	0.99
<i>Language section</i>	0.96
<i>Right-hemisphere section</i>	0.85

Validation: Quantitative analysis

Language Section. One-way ANOVA was performed to determine whether the total Language section of the battery could discriminate between left hemisphere lesion, right hemisphere lesion and control groups. ANOVA's assumption of normality was upheld, but the assumption of homogeneity of variance was violated ($p = 0.000142$). However, ANOVA is a robust technique and sample sizes were equal, so analysis could continue. Table 11 provides descriptive statistics (means and standard deviations) for the analysis. There was a significant main effect, $F(2, 42) = 85.74, p < 0.0000001$. The effect size was determined by calculating *eta*-squared, which was 0.8035. Therefore, group (left hemisphere lesion, right hemisphere lesion and control) accounts for 80.35% of the variation in Language section performances.

Table 11. Mean score for Language section

		Pathology Group		
		Left Hemisphere lesion	Right Hemisphere lesion	Control
Language section score	<i>M</i>	10.20	25.87	29.07
	<i>SD</i>	6.44	3.07	1.62
	<i>N</i>	15	15	15

A post-hoc test (Tukey's HSD) was performed to determine the location of the difference. Correlations between those with left and right hemisphere lesions showed a strong significant difference ($p = 0.000119$). Correlations between left hemisphere lesion group and controls were also highly significant ($p = 0.000119$). Correlations between right hemisphere lesion group and controls were, however, not significant ($p = 0.107411$). Therefore there is significant difference between how left hemisphere lesion group perform against the control and right hemisphere lesion group.

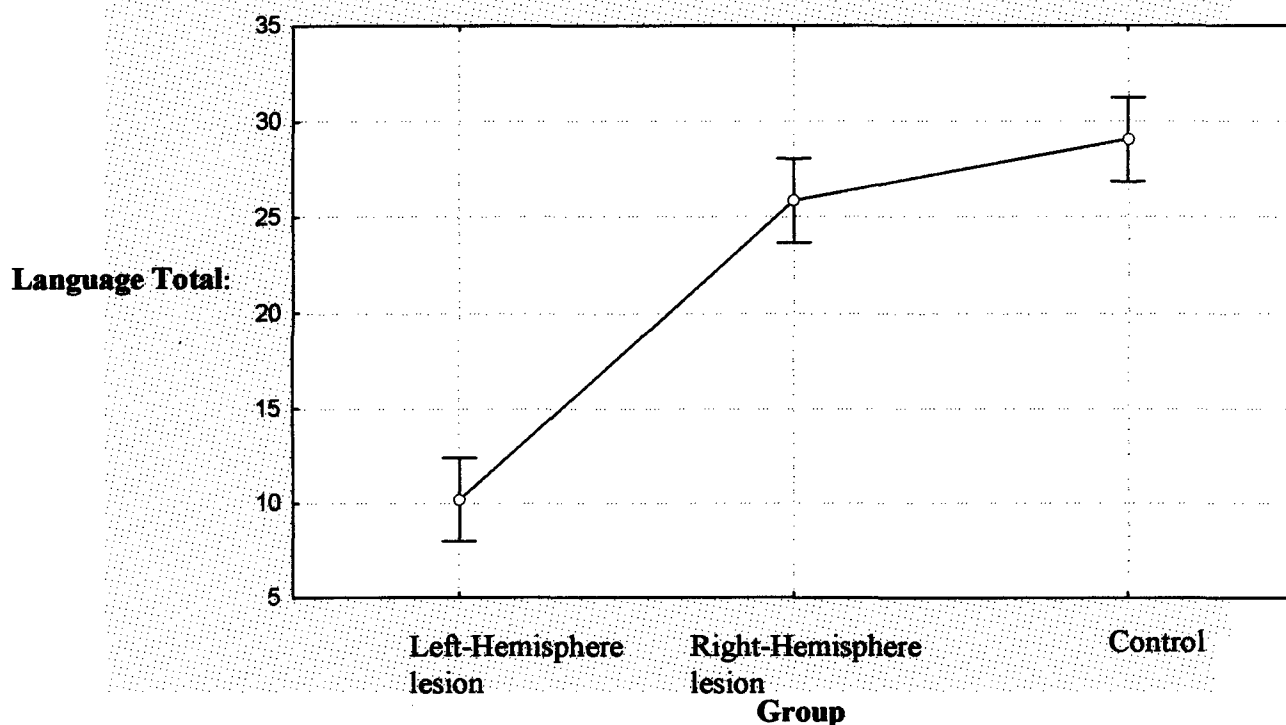


Figure 3. Cell mean plot showing total Language section performances by group

Factorial ANOVA was completed in order to ascertain what factors besides neurological pathology could influence the performance in the total score of the Language section. Assumptions of normality and homogeneity of variance were upheld. Table 12 provides descriptive statistics of analysis. There was no significant main effect for education, $F(1, 39) = 2.11, p = 0.154462$, and no significant effect for language, $F(2, 39) = 0.01, p = 0.988868$. There was no significant interaction effect, $F(2, 39) = 0.94, p = 0.398907$. This shows that education and first language levels did not significantly affect total Language section scores.

Table 12. Mean score for Language section

Factor A: Education	Factor B: Language			A marginal means
	Afrikaans	English	isiXhosa	
8 years or more	¹ 27.33	23.00	22.75	24.32
	² 4.63	10.32	8.54	
	³ 6	9	4	
less than 8 years	¹ 17.50	20.90	21.83	19.81
	² 9.90	9.36	10.63	
	³ 10	10	6	
B marginal means	21.19	21.89	22.20	

¹ Mean ² St. Dev ³ N

A factorial ANOVA investigating the interaction between age and pathology group on the total Language section was also performed. ANOVA's assumption of homogeneity of variance was violated; however, assumption of normality was upheld, and analysis proceeded as pathology groups were equal, as were age groups. Participants were divided into two groups: those younger than fifty years old and those who were fifty years old and older. This division was made in order to carry out the analysis, as ANOVA requires equal sample sizes if the assumption of homogeneity is violated. The fifty-year old cut-off divides the sample into two groups (twenty-three and twenty-two participants).

Table 13 provides descriptive statistics of the analysis. There was a significant main effect for pathology type, $F(2, 39) = 83.52, p < 0.0000001$ and no significant effect for age, $F(1, 39) = 2.87, p = 0.097893$. There was no significant interaction effect, $F(2, 39) = 0.54, p = 0.586940$. Pathology group significantly affected Language section performance, whereas age had no significant influence. This analysis also shows that there was no significant interaction between pathology group and age.

Table 13. Mean score for Language section

Factor A: Age Group	Factor B: Pathology Group			A marginal means
	Left Hemisphere	Right-Hemisphere	Control	
50 years or older	¹ 8.38	25.11	28.60	19.82
	² 6.80	3.69	2.61	
	³ 8	9	5	
younger than 50 years	¹ 12.29	27.00	29.30	23.52
	² 5.77	1.41	0.95	
	³ 7	6	10	
B marginal means	10.20	25.87	29.07	

¹ Mean ² St. Dev ³ N

Naming Test. One-way ANOVA was performed to determine whether the Naming Test could discern left hemisphere lesion participants from controls and those with right hemisphere lesions. Assumptions of ANOVA were followed. Normality was upheld, but homogeneity of variance was violated. As mentioned previously, analysis could continue as sample sizes are equal and ANOVA is robust.

Table 14 shows descriptive statistics for the analysis. There was a significant effect for pathology type, $F(2, 42) = 52.68, p < 0.0000001$. Pathology type greatly influenced Naming Test performance. Tukey's HSD was used to determine where the differences between groups were present. Significant differences were found between the left hemisphere lesion and the right hemisphere lesion groups ($p = 0.000119$) and between the left hemisphere lesion and the control groups ($p = 0.000119$). Effect size was calculated as 0.7419; therefore, 74.19% of the variation in Naming Test performance can be accounted for by group.

Table 14. Mean score for Naming Test

Factor A: Age Group	Factor B: Pathology Group			A marginal means
	Left-Hemisphere	Right-Hemisphere	Control	
50 years or older	¹ 5.38	22.00	25.80	16.82
	² 6.72	2.92	3.27	
	³ 8	9	5	
younger than 50 years	¹ 11.29	23.67	27.30	21.48
	² 8.90	3.08	2.41	
	³ 7	6	10	
B marginal means	8.13	22.67	26.80	

¹ Mean ² St. Dev ³ N

Factorial ANOVA was performed to determine whether factors such as pathology type and age influence Naming Test performance and interact with one another. Age was divided into two groups: those who were younger than fifty years old and those who were fifty years old and older. This was done in order to keep the sample sizes equal. ANOVA's assumption of homogeneity was violated, but assumption of normality was upheld. Analysis continued as sample sizes were equal. Table 14 shows the descriptive statistics for the analysis. There was a significant main effect for pathology type, $F(2, 39) = 50.25, p < 0.0000001$ and no significant effect for age, $F(1, 39) = 1.57, p = 0.217639$. There was no significant interaction effect, $F(2, 39) = 1.57, p = 0.220821$. Therefore, there is no interaction effect between pathology group and age.

A further factorial ANOVA was completed to determine the influence of education and first language on Naming Test performance. ANOVA's assumptions of homogeneity of variance and normality were upheld. Table 15 shows the descriptive statistics (means and standard deviations) for the analysis. There was a significant main effect for education level, $F(1, 39) = 4.11, p = 0.049438$ and no significant effect for language, $F(2, 39) = 0.02, p = 0.984833$. There was no significant interaction effect, $F(2, 39) = 0.58, p = 0.563831$. There is no interaction effect between education level and first language. This shows that the only significant factor influencing Naming Test performance is level of education. Effect size for this analysis was calculated as 0.1228; therefore, 12.28% of variation in Naming Test performance can be explained by education level.

Table 15. Mean score for Naming Test

Factor A: Education	Factor B: Language			A marginal means
	Afrikaans	English	isiXhosa	
less than 8 years	¹ 15.00	18.40	16.67	16.69
	² 9.15	9.80	11.66	
	³ 10	10	6	
8 years or more	¹ 25.00	21.33	22.00	22.63
	² 4.10	11.88	2.94	
	³ 6	9	4	
B marginal means	18.75	19.79	18.80	

¹ Mean ² St. Dev ³ N

Spatial Cognition section. One-way ANOVA was performed to ascertain whether the right hemisphere lesion group performed differently on the total Spatial Cognition section relative to the left hemisphere lesion and control groups. ANOVA's assumptions of normality and homogeneity of variance were both upheld. Table 16 provides descriptive statistics (means and standard deviations) for the analysis. There was a significant effect for pathology type, $F(2, 42) = 29.78, p < 0.0000001$. This shows that there was a significant difference as to how groups performed. Effect size was calculated as 0.7658; therefore, 76.58% of the variation in Spatial Cognition performance can be accounted for by group. Tukey's HSD showed that there was a significant difference between the

performances of the left hemisphere lesion and right hemisphere lesion groups ($p = 0.030567$) and between the control and right hemisphere lesion groups ($p = 0.000119$). Furthermore, there was a significant difference found between the control and left hemisphere lesion groups ($p = 0.000148$).

Table 16. Mean score for Right-Hemisphere section

		Pathology Group		
		Left-Hemisphere	Right-Hemisphere	Control
Spatial Cognition section score	<i>M</i>	7.87	5.13	13.00
	<i>SD</i>	4.00	2.33	1.65
	<i>N</i>	15	15	15

Factorial ANOVA was used to determine whether language and level of education influence performance on the Spatial Cognition section. Assumptions of normality and homogeneity of variance were upheld. Table 17 provides descriptive statistics for this analysis. There was a significant main effect for education level, $F(1, 39) = 4.55$, $p = 0.039257$ and no significant effect for language, $F(2, 39) = 0.69$, $p = 0.506244$. There was no significant interaction effect, $F(2, 39) = 2.10$, $p = 0.135668$. This illustrates that level of education significantly influences performance on the Spatial Cognition section ($p = 0.039257$). Effect size was calculated as 0.2599 for the analysis; therefore, 25.99%

of the variation in Spatial Cognition performance can be accounted for by education level.

Table 17. Mean score for Right-Hemisphere section

Factor A: Education	Factor B: Language			A marginal means
	Afrikaans	English	isiXhosa	
less than 8 years	¹ 7.00	5.70	10.50	7.30
	² 4.27	4.37	1.64	
	³ 10	10	6	
8 years or more	¹ 11.33	10.56	9.25	10.53
	² 3.98	4.07	3.77	
	³ 6	9	4	
B marginal means	8.63	8.00	10.00	

¹ Mean ² St. Dev ³ N

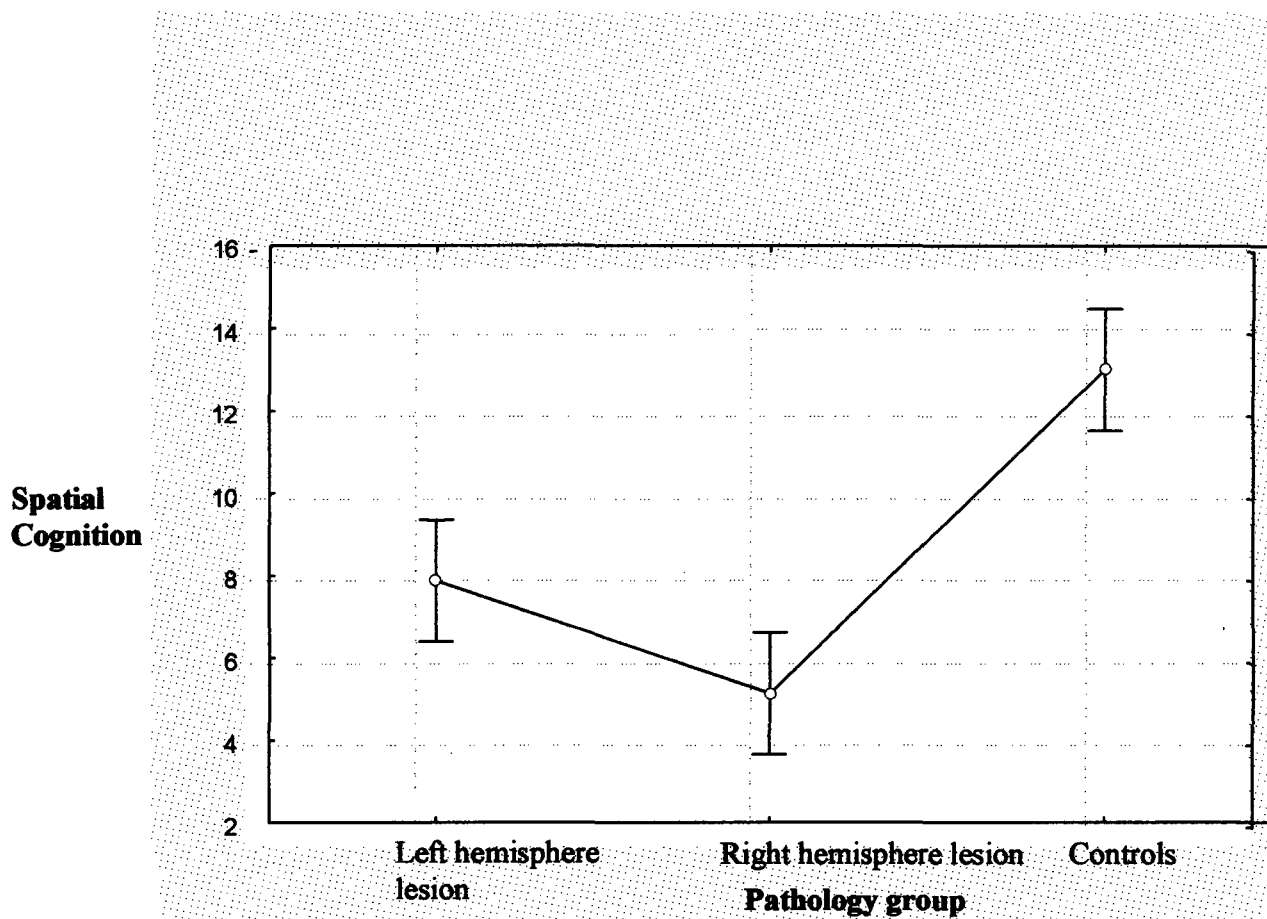


Figure 4. Cell mean plot showing total Spatial Cognition section performances by pathology group

Spatial Acalculia Test. Chi-squared test was performed to ascertain whether the left hemisphere lesion, control and right hemisphere lesion groups performed differently on the 3-D Analysis Test. The Chi-squared statistic was calculated as $\chi^2 = (2, N = 45) = 18.41, p = 0.000101$. Participants' performances on the Spatial Acalculia test are contingent to the group to which they belong (left hemisphere, control or right

hemisphere group). Fourteen participants with right hemisphere lesions failed the Spatial Acalculia test, with only one participant passing the test. Ten controls passed the test with five failing it, and fourteen left hemisphere lesion participants failed the Spatial Acalculia test with only one passing it. A Chi-square test of independence was performed to examine the relationship between pathology group and Spatial Acalculia performance. The relation between these variables was significant, $\chi^2 (1, N = 45) = 1.25, p = 0.263553$. Controls were more likely to perform better on the test than the left hemisphere lesion and right hemisphere lesion group.

A Chi-squared test was performed to ascertain whether control participants with different education levels performed differently on the Spatial Acalculia Test. The Chi-squared statistic was calculated as $\chi^2 = (2, N = 45) = 1.25, p = 0.263553$. This shows that education level was not associated with Spatial Acalculia test performance among controls. However, due to the small sample size in this test, results should be reported tentatively.

Chi-squared test was performed to ascertain whether control participants with different first languages performed differently on the Spatial Acalculia Test. The Chi-squared statistic was calculated as $\chi^2 = (2, N = 45) = 2.4, p = 0.301194$. This shows that the first language of controls was not associated with their Spatial Acalculia test performance. As mentioned earlier, due to the small sample size in this test, these results should be reported cautiously. In order to reduce the control group failure rate (one-third were failing the Spatial Acalculia Test), the cut-off score was lowered from two correct

answers to one. To determine whether this new cut-off score could differentiate between the control and right hemisphere lesion groups, a Chi-squared statistic was performed.

The Chi-squared statistic was calculated as $\chi^2 = (2, N = 45) = 25.79, p = 0.000003$. This shows that Spatial Acalculia Test performances are contingent on participants' group (left hemisphere lesion, control or right hemisphere lesion). The right hemisphere lesion group now have two participants who passed the test. The left hemisphere lesion group's performance remains unchanged with one participant who passed. The control group's pass rate improved, however, from five failures to two.

3-D Analysis Test. A Chi-squared test was performed to ascertain whether the left hemisphere lesion, control and right hemisphere lesion groups performed differently on the 3-D Analysis Test. The Chi-squared statistic was calculated as $\chi^2 = (2, N = 45) = 21.71, p = 0.000019$. Participants' performances on the 3-D Analysis Test are contingent to the group they belong (left hemisphere, control or right hemisphere). All the participants with right hemisphere lesions failed the test, and thirteen from fifteen controls passed the test. However, eleven from fifteen left hemisphere lesion participants failed the test.

A Chi-squared test was performed to ascertain whether left hemisphere lesion participants with different education levels performed differently on the 3-D Analysis Test. The Chi-squared statistic was calculated as $\chi^2 = (1, N = 45) = 0.17, p = 0.679708$.

This shows that education level was not associated with 3-D Analysis test performance among the left hemisphere lesion group.

A Chi-squared test was performed to ascertain whether left hemisphere lesion participants with different first languages performed differently on the 3-D Analysis Test. The Chi-squared statistic was calculated as $\chi^2 = (2, N = 45) = 0.51, p = 0.774388$. This shows that first language was not associated with 3-D Analysis test performance among the left hemisphere lesion group. As mentioned earlier, due to the small sample size in this test, results should be reported tentatively.

Validation: Qualitative analysis

All forty-five scoring sheets that recorded the assessments were given to both researchers, who were still blind to the participants' pathology status. They were required to assign each scoring sheet performance to either the left hemisphere lesion, right hemisphere lesion or control group. Researchers were not provided with the total scores of the screening tool. Diagnoses were made based on performances on tests and qualitative notes made by the assessor.

From the ninety allocations (forty-five per researcher), only one scoring sheet was incorrectly allocated. The incorrectly allocated assessment was from a patient with right hemisphere lesion. The researcher assigned it incorrectly as a left hemisphere deficit script.

6.) Discussion

The goal of this study was to contribute to the eventual provision of a screening tool specifically for South Africa – the Groote Schuur Neurocognitive Battery. This screening tool is being developed to minimise cultural and linguistic bias that would negatively affect South African performances and at the same time provide theoretical underpinnings to the process of screening. The screening tool is required to show validity and reliability, which would increase confidence in its ability to discriminate between neurocognitively healthy controls and neurocognitively impaired patients and between patients with left hemisphere lesions versus right hemisphere lesions on the basis of the distinct neurocognitive pictures typical of each.

The first phase of this study was to finalise adaptations of two tests used in the Neurocognitive Battery. Two years ago, once new tests identified for inclusion in the battery had been initially adapted and piloted, the Naming Test and the 3-D Analysis Test were isolated as still being problematic. Control participants were performing too well on the Naming Test, which was causing a slight ceiling effect. More difficult items were needed that were also culturally fair in order not to discriminate between the three language groups (English, Afrikaans and isiXhosa). The goal of the re-piloting phase was to minimise the influence of participants' first language on the tests and to improve controls' performance on the tests. The 3-D Analysis Test, prior to the re-piloting carried out in this present study, had a distorted perspective and had elicited poor performances from controls. The re-piloting phase's goal was therefore to pilot the new 3-D Analysis Test with corrected block configurations and explore reducing the difficulty of the test.

Re-piloting phase

Naming Test

Results showed that performances on the four possible new Naming Test items were not influenced by participants' first language. The two items that needed to be replaced had previously shown distinct language bias, with isiXhosa-speaking participants performing more poorly than their Afrikaans and English-speaking counterparts.

Qualitative observations confirmed that the items "aloe" and "pylon" were the most appropriate two of the four to be included in the final test. While they were difficult items and would help reducing the ceiling effect of the Naming Test, they were also known by the majority of the three language groups. Results demonstrated that the new Naming Test with the "hippopotamus" and "mug" items replaced by "aloe" and "pylon" produced similar averages among the language groups, which indicates that they are performing alike. Furthermore, all standard deviations among the language groups were also comparable (ranging from 3.05 to 3.46), which indicates that these groups' performances were varying comparably.

Furthermore, it was investigated whether 'level of education' influenced Naming Test performance and whether it interacted with the participants' first language. Results showed that while the participants' level of education significantly affected their Naming Test performance, first language did not and that there was no interaction between level of education and first language. The effect of the participants' level of education on

Naming Test performance was shown to be very significant ($p = 0.002345$), where those with less than eight years of education performed much worse on the Naming Test than those with eight years of education or more. It is well documented that education level has a large effect on test performance (Ostrosky et al. 1998), where those with lower levels of education perform significantly worse on neuropsychological tests than those with higher levels of education. Performances on visual confrontation naming tests are particularly sensitive to education levels (Mansur et al. 2005; Talberg, 2005).

It is crucial that neuropsychological tests are developed to eliminate language bias to the groups they were intended. Swartz, Drennan and Crawford (1997) and Swartz (1998) admonish against the profound effect language has on the assessment of patients and Rosselli and Ardila (2003) warn that language tests are especially susceptible to language and cultural bias. To have a language test show that it is not influenced by a participant's first-language (English, Afrikaans and isiXhosa) has enormous practical value. The tester can interpret the Naming Test performance of his or her patient immediately without being wary of incorrectly diagnosing his or her patient with aphasia due to language bias.

It was also investigated how age influenced performance on the Naming Test. There was a significant difference between how participants who were older than forty years old performed against those who were forty years old and below ($p = 0.013508$). The older half of the sample performed significantly more poorly than the younger half of the group. Carter et al. (2005) maintains that age can have a significant effect on neuropsychological test performance. However, age and quality of education may have

an interaction effect on neuropsychological performances. The older, black generation of South Africans were forced into separate and inferior education systems (*Bantu education*) during the Apartheid regime. The effects of education and age on the Naming Test suggest the need for separate tests for lower education and older populations, or at least, examiners who are aware of the effect these factors have on test performances.

The main goal of the re-piloting of the Naming Test was to construct a test that had a lower overall average total score to avoid a ceiling effect, and one that produced similar performances among language groups. This goal has been achieved. Of interest, the mean of the new Naming Test was 25.97 with a standard deviation of 3.13. The United States norm for the Boston Naming Test is 25.4 with a standard deviation of 3.1 (Williams et al. 1989 in Lansing, Ivnik, Cullum and Randolph, 1999). While not significant in that this study is not comparing South African performances to United States norms, it is encouraging that the ceiling effect of the Naming Test has been reduced because the distribution of scores takes on a similar shape.

Re-piloting the 3-D Analysis Test

Comparisons between the originally adapted Cube Analysis Test and the latest 3-D Analysis Test further modified in this study showed no significant difference. This reveals that control participants were not performing better on the new 3-D Analysis Test after the perspectives and size of the blocks were corrected and enlarged. Further statistical analyses were done to learn whether the poor performances were the result of the participants' first languages or education levels. Results showed that education was a

huge influence on 3-D Analysis Test performance, with the $p = 0.000482$ illustrating just how significant the 'level of education' variable was on 3-D Analysis Test performance. Rosselli and Ardila (2003) argue that visuo-spatial tests are school-dependant abilities and that level of education has considerable influence on verbal and non-verbal test performances.

There was a slightly significant interaction effect between the re-piloting group's education level and first-language ($p = 0.04845$) on the 3-D Analysis Test. Cell mean plots showed the interaction effect was present at the isiXhosa-language level. The performances of the isiXhosa-speaking participants were unaffected by their education level (greater and less than eight years of education); that is, their performances on the 3-D Analysis Test were similar regardless of their education level. English-speaking participants were the most susceptible to education level. English-speaking participants with a higher level of education performed better than those with a lower level of education. However, the significance of this finding is questioned, as the probability value obtained (0.045845) was very close to the cut-off probability of 0.05. Future research work conducted on the Neurocognitive Battery and its tests should investigate whether there is indeed a clear interaction effect between language and education on 3-D Analysis Test performance.

Age did not influence performance on the 3-D Analysis Test. However, Carter et al. (2005) argue that age can have a significant effect on neuropsychological performance. The average age of the sample for this part of the study was 44.22 years old with a

standard deviation of approximately fifteen years, which shows that a relatively young sample was used. A difference in performance due to age would be more overt with a sample that has a larger variation in ages and an older average.

The 3-D Analysis Test performances were quite poor, with an average of 8.63 from 14. The cut-off score in the original battery was 10. In order to make it easier to raise the controls' scores, frequencies of individual items' scores were created. This was performed to ascertain which items elicited the weaker performances. From the six most poorly-answered items, four were removed from the test to produce a newer version of the 3-D Analysis Test, comprising ten items. The new average was calculated as 7.03 from 10, which was higher than the previous version (6.16 from 10). Further analyses showed that this new version was not influenced by language or age, but was influenced by education level. These results are similar to those of the original 3-D Analysis Test.

In concluding this section, the re-piloting section was a success in that the Naming Test average among controls was lowered to reduce a ceiling effect, and it was also ascertained that language did not influence test performance. However, education level did influence the performance, which is compatible with previous literature which maintains that education can considerably influence neuropsychological performance (Ostrosky et al. 1998).

The latest version of the 3-D Analysis Test, however, was shown to be similar to its predecessor, the originally adapted 3-D Analysis Test. Items were removed from the test

in an attempt to improve the control group's performance. It was also shown that level of education significantly influenced performance on the 3-D Analysis Test, which coincides with Rosselli and Ardila's (2003) argument that education level influences visuo-spatial tests. Nonetheless, the updated ten-item 3-D Analysis Test was deemed suitable for inclusion in the final Neurocognitive Battery ahead of its validation, as was the latest Naming Test.

Reliability

The second phase of this study was to establish the reliability of the Language and Spatial Cognition sections. These sections have to show consistency over repeated assessments, and across examiners (Durrheim, 1999). The correlation coefficients for all individual sections (Orientation, Language function and Spatial Cognition function) were very high, ranging from 0.84785 to 0.99398, with an overall correlation coefficient of 0.89458. According to Guilford's interpretation of correlation coefficients, any coefficient larger than 0.7 shows a high correlation with a strong relationship and any coefficient larger than 0.9 shows a very high correlation with a very dependable relationship (Sprinthall, 1987 in Lachenicht, 2002). This shows that the Language and Spatial Cognition sections display very strong reliability – consistency over assessment and across examiners. These sections of the battery are therefore dependable.

Lezak (1995) warns that it can be difficult to establish reliability for neuropsychological assessments, as neuropsychological impairments can vary from day to day or resolve quickly. Fan et al. (1988) in Lezak (1995) argue that visuo-spatial inattention is

commonly known to vary. This might explain why the Spatial Cognition section of the battery produced the lowest correlation coefficient ($r = 0.84785$), which was still however, very strong.

Validation phase: Quantitative

Validity needed to be established to determine both statistically and using qualitative feedback that the two sections of the Neurocognitive Battery were able to discern accurately between three groups: neurocognitively intact individuals (controls), participants with left hemisphere lesions, and participants with right hemisphere lesions. It was also vital to establish what other variables were influencing performances on the two sections, such as education level and first language. The final aspect of confirming validation was to have two researchers evaluate the performances of the participants and to 'allocate blindly' to which group each performance belonged, while not having access to the final scores but rather basing their decision solely on the qualitative test performance, in keeping with a Lurian approach (Luria, 1999).

Language section:

The central goal for providing validity for the Language section was to show that the left hemisphere lesion groups' performances on the Language section successfully differentiated them from both the control and right hemisphere lesion groups on the same section. The results overwhelmingly showed that the left hemisphere lesion group performed much worse than the control and right hemisphere lesion groups, which highlights the Language section's ability to differentiate between those with language

dysfunctions and those without. To illustrate the enormity of the influence that group had on performance on the Language section: the *eta*-squared, or the effect size, showed that 80.35% of the variation in test performance in the Language section could be explained by the participants' group.

Further analyses were performed to investigate whether language, education level and age affected Language section performance. None of the variables significantly influenced the participants' performances on the Language section. Examiners can interpret a patient's performance on the Language section without uncertainty regarding first-language. However, while age and education levels were insignificant in this analysis, education level was only analysed between those with less and more than eight years of education. Further studies should be made to ascertain the difference in performance across a range of education levels. Similarly, this sample was divided into two groups for age (those older and younger than forty-five years old). A similar study comparing a much older group with younger groups would be beneficial in ascertaining the influence age has on the Language tests.

However, these results are encouraging in showing the reduction in bias of the tests. It strengthens confidence in an assessor that poor performances on these language tests are indicative of true pathology and not clouded by language bias.

Investigations were carried out to evaluate the efficacy of the new Naming Test. It was shown to be able to differentiate between left hemisphere lesion participants and controls

and left hemisphere lesion and the right hemisphere lesion groups. Probability values were very small ($p = 0.000119$) which shows that there was a highly significant difference regarding how participants with left hemisphere lesions performed compared to their control and right hemisphere lesion counterparts. Further analyses showed that the test was not influenced by age or first language spoken, but was influenced by education level. This is a common finding, in that many visual confrontation naming tests are sensitive to education bias (Kennepohl, et al., 2004; Manly et al., 2004; Mansur et al., 2005).

However, while there was a significant difference between how the two education groups performed, the probability value was calculated as 0.049438. This is very close to the significance level cut-off of 0.05. While this indicates that there is a significant difference between education groups, both groups were within acceptable limits of normal functioning. Therefore, analyses for the Naming Test show that it is not overtly affected by language, age or education level. However, further studies with varying age and education levels could confirm the impact of age on education level on Naming Test performance.

Spatial Cognition section

The primary goal for the validation of the Spatial Cognition section of the Neurocognitive Battery was to determine whether the section could successfully discriminate between right hemisphere lesion participants versus participants with left hemisphere lesion participants; and between participants with right hemisphere lesions

versus control performances. Results showed that the Spatial Cognition section of the battery was able to discriminate between those with right hemisphere lesion and controls ($p = 0.000119$); and between those with left hemisphere lesion and right hemisphere lesion ($p = 0.030567$). There was a large difference in how the groups performed in the Spatial Cognition section. There was also a significant difference between how left hemisphere lesion performed compared to controls ($p = 0.000148$). This can be understood in that patients with language dysfunction will perform poorly in many neuropsychological tests that do not specifically assess language. Difficulties comprehending speech can potentially influence any test, especially when language is the primary medium through which assessments are performed (Swartz, 1998).

Furthermore, spatial cognition is not an exclusively right hemisphere function. For example, left parietal lesions can lead to spatial impairments which are qualitatively different to impairments caused by right hemisphere lesions (Devinsky & D'Esposito, 2003). Therefore, the qualitative performances of the patient need to be taken into account in order to discern whether he or she failed the Spatial Cognition section because of a left- or right-hemisphere deficit of spatial cognition. For example, a patient with a left hemisphere lesion may fail the Rey Complex Figure copy in the Groote Schuur Neurocognitive Assessment Battery because there is detail lacking in the picture while the overall gestalt of the figure is intact (Devinsky & D'Esposito, 2003). A patient with a right hemisphere lesion could fail the Rey Complex Figure because the overall shape of the figure is distorted and/or left-sided neglect is present. The battery should explicitly provide guidelines on distinguishing left hemisphere lesion performances from right

hemisphere lesion performances. A more thorough multi-step approach to assessing performance in the Spatial Cognition section is needed. For example, the scoring of the Rey Complex Figure in the Groote Schuur Neurocognitive Battery should include multiple steps where a typical left-hemisphere lesion performance is eliminated before ascertaining whether the performance is more indicative of a right hemisphere lesion. This decision tree approach would be able to clearly differentiate between left- and right-hemisphere lesions.

Further analyses demonstrated that the participants' first language did not influence their performance on the Spatial Cognition section. This is congruous with previous analyses that have shown that all tests mentioned were not influenced by language level. However, education level did influence performance on the Spatial Cognition section ($p = 0.039257$). Participants with less than eight years of education performed worse on the Spatial Cognition section than those with more than eight years of education. This is in agreement with Rosselli and Ardila (2003), who maintain that visuo-spatial tests are affected by education level. Interpretations regarding a patient's performance on these tests should be made while considering the impact of education.

3-D Analysis Test

Results on the 3-D Analysis Test showed that there was a significant difference between control and right hemisphere lesion group performances. Only two control participants failed the test, and none of the right hemisphere lesion patients passed. However, the left hemisphere lesion group performed poorly on the test in comparison, with only four

participants passing the test. As mentioned earlier, participants with severe language impairment would be at a disadvantage regarding most tests. The 3-D Analysis Test is also a particularly difficult visuo-spatial test (Lezak, 1995), and during assessments, it was noted that most participants (even controls) needed a thorough explanation during the three-practice-items component of the test. An aphasic with even mild comprehension difficulties would not have the advantage of the practice items. A recommendation for the overarching study would be to simplify this test or to replace it with an easier one, as there appears to be a floor effect present.

Further analysis showed that education level and first language did not influence the participants' performances on the 3-D Analysis Test. This is an encouraging finding, in that interpretations of 3-D Analysis Test performance can be made without overt interference by education level and first language. Again, this is an important finding because it suggests that interpretations can be made regarding a patient's 3-D Analysis Test performance without being unable to decide whether the patient had a deficit or whether the poor performance was the result of education level and first language influencing the test performance.

Another sub-test, the Spatial Acalculia test, was investigated from the Spatial Cognition section because of its inability to discern between left hemisphere lesion and right hemisphere lesion participants. Furthermore, control participants performed poorly on this test.

Spatial Acalculia test

Spatial Acalculia test results revealed that there was a significant difference between how controls, left- and right-hemisphere lesion participants performed ($p = 0.000101$). Controls performed differently to the right hemisphere lesion group but there was no difference in test performance between the left- and right-hemisphere lesion groups. Therefore the test is able to discriminate between the control and right hemisphere lesion groups; but not between the left- and right-hemisphere lesion groups. While there was a significant difference between control and right hemisphere lesion groups, one-third (five) of the controls still failed the test.

Further analyses showed that performance on the Spatial Acalculia test was not influenced by education or first language, together with the previously mentioned tests contained in the Spatial Cognition section. However, the high control group failure rate (five failures from fifteen participants) was concerning. The cut-off score of two or more correct calculations was then reduced to one or more correct calculations.

It can be argued that the arithmetic calculations are too difficult and should be simplified, or the cut-off scores should be changed. Initially, the cut-off score was two, where anyone who only performed one calculation correctly or none would fail the test. Further analyses were made to observe whether the controls would fair better if the cut-off score were reduced from two calculations correct, to only one.

The results, with the new cut-off score implemented, showed that the test could still differentiate between participants with right hemisphere lesion and controls. However, now, thirteen of the fifteen controls were passing the test. A recommendation would be to retain the new cut-off score to make the calculations easier for neurocognitively intact patients to perform well, or to introduce easier calculations.

A further concern regarding the Spatial Acalculia test results was the poor performances from the left hemisphere lesion group. Ardila and Rosselli (2002) propose that calculation ability is not only dependent on spatial ability, but verbal as well, amongst others. The test is used in the battery to ascertain spatial acalculia, a result from visuo-spatial defects which can impact the spatial organisation of the sums. However, those who cannot comprehend the arithmetic operators, such as aphasics, would also be unable to perform calculations. They further argue that many aphasics would present with primary acalculia, which is an inability of computational ability, or anarithmetia (Ardila & Rosselli, 2002). It can be argued that this is not a suitable test to discern between those with spatial impairment and language deficit.

However, the quality of the performance is crucial in discriminating between the left- and right-hemisphere lesion groups, and it should be implemented into the decision tree organisation of the test. Ardila and Rosselli (2002) argue that in Broca's aphasics, patients have shown a deficit in the syntax of the calculation, where "These patients present "stack" errors (e.g., 14 is read as 4) that could be interpreted as an agrammatism

in the numerical system. They also have difficulties counting backward and in successive operations (e.g., 1, 4, 7, or 20, 17, 14)” (Ardila & Rosselli, 2002:203).

Wernicke’s aphasics show dysfunction in lexical processing: their comprehension deficit prevents them from reading basic arithmetical signs. These patients can also have problems retaining the rules and progress of the calculation from a verbal memory deficit (what Luria understood as acoustic-amnesia, a sub-type of Wernicke’s aphasia) (Luria, 1973 in Ardila & Rosselli, 2002).

Spatial acalculia presents with poorer written calculation than mental calculation (Ardila & Rosselli, 2002). Patients with spatial neglect may also ignore the digits on the left side. The spatially impaired patient, while understanding the operators in the calculations, will have difficulties knowing where in columns the numbers should be placed (Ardila & Rosselli, 2002). The Spatial Acalculia test is an appropriate example of how this Neurocognitive Battery should be used. Emphasis is placed on qualitative analysis where the decision tree structure aids in understanding not only that the patient passed or failed the test, but why he or she did and to which area of the brain the deficit is related.

The Spatial Acalculia test in the Groote Schuur Neurocognitive Assessment Battery uses a two-step, qualitative approach. Firstly, the participant is provided with simple arithmetic calculations verbally to assess Primary Acalculia and then moves on to more complicated calculations to assess Spatial Acalculia. However, the Neurocognitive Battery does not provide enough detail to properly establish Primary Acalculia.

Specifically, simple sums need to be provided in order to prevent spatial-loading from taking place. Spatial-loading refers to the fact that mental arithmetic can draw on cognitive spatial ability once the task requires the person to move numbers over decimal places. However, simple verbal calculations do not require spatial ability. These simple calculations must not involve moving numbers over decimal places (for example, $5 + 7$), but should only ascertain primary acalculia (for example, $5 + 3$). If the patient is able to complete these verbal calculations correctly, they may move on to the spatial acalculia questions. Having done this, if the patient then goes on to fail the spatial acalculia questions, the assessor can be confident that it is not due to primary acalculia as this option has already been eliminated.

Validation phase: Qualitative

From the allocations of the participants to each of the three participant groups carried out by the two researchers, each reviewing the forty-five participants independently, only one scoring sheet was incorrectly allocated. The incorrectly allocated assessment was from a patient with a right hemisphere lesion. The researcher assigned it incorrectly as a left hemisphere lesion participant's script. The findings are encouraging in that 98.89% of allocations were correct. Allocations were based on test performances and qualitative notes; scores of the sections were excluded.

Language section

Both researchers found the Language section's performances straightforward to interpret. Right hemisphere lesion participants and controls performed fluently on the Washing

Line Picture Test. A few of the right hemisphere lesion participants, suffering from unilateral neglect, did not discuss activities that were occurring on the left side of the picture. Qualitative notes in the scoring sheet drew attention to that feature. Most of the right hemisphere lesion participants and controls performed well on the Naming Test.

Aphasics would commonly display literal and semantic paraphasias or display word-finding difficulties in the Naming Test. One aphasic participant incorrectly named the 'carrot' a 'carriage' and when shown a picture of a 'spear' called it a 'spade'. A spatially impaired participant who also performed poorly on the Naming Test thought that a 'soccer net' was 'windows', that a 'pot' was a 'bird' and that the 'grasshopper' was a 'rowing boat'. The latter participant's performance does not indicate word-finding difficulty or paraphasias, but rather a spatial perceptive impairment.

Education level did impact on the control group's performance. While performing the writing test, controls with low education levels could not spell correctly, even when asked to write any spontaneous sentence (common mistakes involved replacing "their" with "there"); however, they did perform well on fluency, comprehension and repetition, which made allocating their scoring sheet correctly easier.

Controls did not show difficulty with the repetition section; however, many of the aphasics could not repeat the complex sentences that the control group could. However, two patients with spatial impairment performed quite poorly on the repetition section as they refused to follow instructions. Instead of repeating "this doctor does not visit all the

patients in the ward”, both participants would repeatedly make derogatory comments about their doctor. An aphasic who failed the test, when asked to repeat “the painter painted many beautiful scenes”, responded with, “the painter painted meeneebetti...”, showing an actual dysfunction in repetition. While the right hemisphere lesion participants performed poorly on the test, they did not display an aphasic performance but rather emotional inappropriateness, which is associated with right hemisphere lesions (Devinsky & D’Esposito, 2003).

Spatial Cognition section

Participants from the left hemisphere lesion group performed quite poorly on most tests in this section, depending on the severity of their dysfunction, which made allocations between the left- and right-hemisphere lesion groups more challenging. However, comments in the scoring sheet indicated that they showed difficulties in understanding the instructions related to the 3-D Analysis Test. A few of the Broca’s patients would perseverate when undertaking the test by repeatedly counting every configuration as containing the same number of blocks. Many Broca’s patients illustrate perseveration due to left frontal lobe involvement. Devinsky and D’Esposito (2003) argue that perseveration can be misunderstood as visuo-spatial dysfunction. However, even though the left hemisphere lesion participants’ performances on the Spatial Cognition section would be poor, their performance on the Language section would be as weak or worse, which indicated that their dysfunction stemmed from a disruption in the language area of the brain that was affecting other functions.

Similar qualitative observations were made for the Spatial Acalculia test. Spatially impaired participants would show problems with the Spatial Acalculia test by incorrectly summing columns, and those with unilateral neglect would ignore numbers on the left side. This performance was often coupled with constructional apraxia and/or neglect on the Rey Complex Figure and the Hut Drawing Test.

Some spatially impaired participants would show emotional inappropriateness throughout the battery, especially with tests they found challenging (the Spatial Cognition section). During the Rey Complex Figure or 3-D Analysis Test, some spatially impaired patients would become easily distracted, would refuse to make eye contact with the assessor and would need encouragement and prompting in order to resume the tests. Only one left hemisphere lesion participant showed this behaviour, whilst none of the controls did.

Some controls would perform badly on the Spatial Acalculia test, but while their final answers were incorrect, their approach was not. They understood the arithmetic operators, did not have a problem lining up numbers in columns and did not ignore any numbers. Importantly, they performed well in the rest of the Spatial Cognition section. Allocating control performances became straightforward through observing how they failed in order to discriminate between a control, left hemisphere lesion or right hemisphere lesion fail.

Qualitative and quantitative analysis

It was important to attain validation of the Neurocognitive Battery both qualitatively and quantitatively, since its development was based on the theoretical foundation of qualitative assessment and the practical application of quantitative methods. Qualitative assessment is imperative in discovering the exact nature of the neuropsychological deficit and how it affects other brain functioning, thus leading to an understanding of the brain areas causing the disruption and the mechanism that led to the dysfunctional behaviour (Akhutina & Tsvetkova, 1983). Quantitative results are able to determine the degree of impairment (Lezak, 1995).

Akhutina and Tsvetkova (1983) warn that a previous attempt at marrying psychometric methods with Lurian theory (the Luria-Nebraska test battery) resulted in a scoring system that relied only on the quantitative nature of a performance (whether a patient passed or failed) and ignored the qualitative aspect (how the patient passed or failed).

They provide an example in the Luria-Nebraska test battery, where the scoring procedures merely state that a patient passed or failed a motor action. The Luria-Nebraska test battery ignores the nature of the performance – whether perseveration or echopraxia was present, for example (Akhutina & Tsvetkova, 1983). It is crucial that the Groote Schuur Neurocognitive Battery incorporate qualitative features into all the decision-trees in the battery itself.

Limitations

Varying education and age levels

Education level and age were divided into only two groups due to the relatively small sample size. Further studies should be initiated to explore what effect other divisions of age and education level would have on the Neurocognitive Battery performance. Groups should include those with very low years of education compared to those with primary/junior school education and secondary/ high school education. Our country has an ignoble history where different racial groups were segregated and black citizens were given an inferior education (*Bantu* education). Because of this, further studies should also include the quality of education participants experienced.

The impact of age on test performance should also be explored with an older group used than the one in this study. However, due to the relatively small sample size and time constraints, such an investigation was impossible.

Level of acculturation

Coffey, Marmola, Schock and Adamsa (2005) argue that level of acculturation is a factor that may influence performance on neuropsychological tests. They refer to acculturation as changes that result when people from other cultures are in constant contact with one another. They warn that level of acculturation is hardly ever reported in neuropsychological research (O'Bryant, O'Jile & McCaffrey, 2004 in Coffey et al., 2005). This study did not explore the impact of acculturation on the tests' performances. Further explorations regarding the impact of this variable would further understanding on

influencers of test performance. Coffey et al. (2005) point out that the length of one's education and location is an indication of one's level of acculturation. A further study should also incorporate participants' place of residence (urban or rural areas) as a variable.

Recommendations and future directions

Spatial Acalculia test

The recommendations for the overarching study, and for possible future studies, are firstly to provide simpler calculations in the Spatial Acalculia test. This is needed as there appears to be a floor effect present as a third of the controls failed the test. Instead of the lowered cut-off scores at the moment, which reduce the scoring of the test to a pass or fail, simpler calculations would allow the battery to differentiate between varying levels of spatial acalculia. At the moment, the Spatial Acalculia test comprises of very simple verbal arithmetic sums and quite difficult calculations (for example, $278 + 843$). A range of easy to difficult spatial acalculia questions should be provided to ascertain the degree of functioning in the participant. This section should be researched further to establish appropriate cut off scores to determine when spatial acalculia is present.

As argued earlier, the 'decision-trees' should incorporate more multi-step guidelines to differentiate between spatial acalculia and other acalculias (such as aphasic acalculia). Scores should not be solely based on whether the calculations were correctly answered but also on how they were approached. In the Language section, for example, the scores

take note of paraphasias and comparisons between repetition and fluency. However, no such qualitative aspect is involved with the Spatial Acalculia test. For example, it is important that the scoring section of the Spatial Acalculia test incorporate information regarding Primary Acalculia performance. If the participant failed the Spatial Acalculia test, Primary Acalculia can be eliminated as a factor.

Influence of age, education and quality of education

Further studies should be done to ascertain the exact range of influence age and education level has on the Neurocognitive Battery's tests. Larger sample sizes with varying levels of age and education level would be helpful in further investigations. Furthermore, there is reason to suspect that quality of education, notably, the *Bantu* education system during Apartheid, could be a factor on test performance (Agranovich & Puente, 2007). A study investigating the influence of quality of education would be useful.

The next phase for this study involves confirming weightings of the final tests. In the scoring procedure, each test constitutes a particular share of the total section's score. Tests that have shown to be most diagnostically useful should constitute a larger share of the total score. For example, the 3-D Analysis and Spatial Acalculia tests have elicited poor performances from the left-hemisphere lesion group. Tests that are able to identify spatial impairment and easily discriminate between groups should have the largest weighting in the Spatial Cognition section. For example, the Rey Complex Figure and Hut Drawing test may be useful as the principal Spatial Cognition tests.

7.) Conclusion

This study has attempted to facilitate in the provision of the first South African-specific neuropsychological screening tool. The implications of such an instrument are important. There is a huge need for a South African-specific screening tool (Wind et al., 1997). Previously, the Mini-Mental State Examination (MMSE) has been used as the sole neurocognitive screening tool at Groote Schuur Hospital. The MMSE has shown to be inadequate as a general screening tool (Faustman et al., 1990; Grace et al., 1995; Nysa et al., 2005) and inappropriate when applied to diverse cultural groups (Stewart et al., 2002).

This study has attempted to provide a more suitable alternative by adapting neuropsychological tests for the South African population. These adaptations were made to reduce language bias and provide tests that neurocognitively intact controls would perform normally. The study has shown that these tests are able to differentiate between control and pathology groups and show no significant language bias. This provides a sense of confidence in the examiner, as a patient's performance can be interpreted without fear of language bias impacting on his or her performance.

The validity and reliability phase has successfully provided confirmation that the Neurocognitive Battery can differentiate between left and right hemisphere lesions. Furthermore, strong inter-rater and test-retest reliability has been demonstrated. This contribution to South African neuropsychology is great in providing the first steps in achieving a theory-driven screening tool especially developed for South Africans.

8.) References

Agranovich, A.V. & Puente, A.E. (2007). Do Russian and American normal adults perform similarly on neuropsychological tests? Preliminary findings on the relationship between culture and test performance. *Archives of Clinical Neuropsychology*, 22, 273-282.

Akhutina, T.V. & Tsvetkova, L.S. (1983). Comments on a Standardized version of Luria's tests. *Brain and Cognition*, 2, 129 -134.

Ardila, A. (2007). Toward the development of a cross-linguistic naming test. *Archives of Clinical Neuropsychology*, 22, 297-307.

Ardila, A. & Rosselli, M. (2002). Acalculia and dyscalculia. *Neuropsychology Review*, 12, 179-231.

Azouvi, P., Samuel, C., Louis-Dreyfus, A., Bernati, T., Bartolomeo, P., Beis, J.M., Chokron, S., Leclercq, M., Marchal, F., Martin, Y., de Montety, G., Olivier, S., Perennou, D., Pradat-Diehl, P., Prairial, C., Rode, G., Sieroff, E., Wiart, L. & Rousseaux, M. (2002). Sensitivity of clinical and behavioural tests of spatial neglect after right hemisphere stroke. *Journal of Neurology, Neurosurgery and Psychiatry*, 73, 160-166.

- Barker-Collo, S. L. (2001). The 60-item Boston Naming Test: Cultural bias and possible adaptations for New Zealand. *Aphasiology*, *15*, 85 – 92.
- Barker-Collo, S. & Feigin, V. (2006). The impact of neuropsychological deficits on functional stroke outcomes. *Neuropsychological Review*, *16*, 53-64.
- Brauer Boonea, K., Victor, T.L., Wen, J., Razani, J. & Pont'on, 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.
- Byrd, D.A., Sanchez, D. & Manly, J.J. (2005). Neuropsychological test performance among Caribbean-born and U.S.-born African American elderly: The role of age, education and reading level. *Journal of Clinical and Experimental Neuropsychology*, *27*, 1056-1069.
- Carter, J.A., Lees, J.A., Murira, G.M. & Gona, J., Neville, B.G.R. & Newton, C.R.J.C. (2005). Issues in the development of cross-cultural assessments of speech and language for children. *International Journal of Language & Communication Disorders*, *40*, 385-401.
- Christensen, A. & Caetano, C. (1999). Luria's neuropsychological evaluation in the Nordic countries. *Neuropsychology Review*, *9*, 71-78.

- Coffey, D.M., Marmola, L., Schock, L. & Adamsa, W. (2005). The influence of acculturation on the Wisconsin Card Sorting Test by Mexican Americans. *Archives of Clinical Neuropsychology*, 20, 795-803.
- 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.
- DeÂmonet, J. & Thierry, G. (2001) Language and Brain: What is up? What is coming up? *Journal of Clinical and Experimental Neuropsychology*, 23, 49-73.
- Devinsky, O. & D'Esposito, M. (2003). *Neurology of Cognitive and Behavioral Disorders*. New York: Oxford University Press.
- Durrheim, K. (1999). Quantitative measurement. In M. Terre Blanche & K. Durrheim (Eds.), *Research in practice Applied methods for the social sciences*. (pp. 72 – 95). Cape Town: UCT Press.
- Durrheim, K. & Wassenaar, D. (1999). Putting research into practice: writing and evaluating research proposals. In M. Terre Blanche & K. Durrheim (Eds.), *Research in practice Applied methods for the social sciences*. (pp. 54 – 71). Cape Town: UCT Press.

Fastenau, P.S., Denburg, N.L. & Mauer, B.A. (1998). Parallel short forms of the Boston Naming Test: Psychometric properties and norms for older adults. *Journal of Clinical and Experimental Neuropsychology*, 20, 828-834.

Faustman, W.O., Moses Jr, J.A. & Csernansky, J.G. (1990). Limitations of the Mini-Mental State Examination in predicting neuropsychological functioning in a psychiatric sample. *Acta Psychiatrica Scandinavica*, 81, 126-131. Abstract retrieved November 26, 2007, from Blackwell-Synergy Database.

Finton, M.J., Lucas, J.A., Graff-Radford, N.R. & Uitti, R.J. (1998). Analysis of visuospatial errors in patients with Alzheimer's disease or Parkinson's disease. *Journal of Clinical and Experimental Neuropsychology*, 20, 186-193.

Fortuny, L.A.I., Garolera, M., Romo, D.H., Feldman, E., Barillas, H.F., Keefe, R., Lemaitre, M.J., Martin, A.O., Mirsky, A., Monguio, I., Morote, G., Parchment, S.A., Parchment, L.J., Da Pena, E., Politis, D.G., Sedo, M.A., Taussik, I., Valdivia, F., De Valdivia, L.E. & Maestre, K.V. (2005). Research with Spanish-speaking populations in the United States: Lost in the translation a commentary and a plea. *Journal of Clinical and Experimental Neuropsychology*, 27, 555-564.

- Gasquoine, P.G. (1999). Variables moderating cultural and ethnic difference in neuropsychological assessment: The case of Hispanic Americans. *The Clinical Neuropsychologist*, 13, 376 – 383.
- Gasquoine, P.G. (2001). Research in clinical Neuropsychology with Hispanic American Participants: A review. *The Clinical Neuropsychologist*, 15, 2 – 12.
- Giannakou, M. & Kosmidis, M.H. (2006). Cultural appropriateness of the Hooper Visual Organization Test? Greek Normative Data. *Journal of Clinical and Experimental Neuropsychology*, 28, 1023-1029.
- Glozman, J.M. (1999). Quantitative and qualitative integration of Lurian procedures. *Neuropsychology Review*, 9, 23-32.
- Grace, J., Nadler, J.D., White, D.A., Guilmette, T.J., Giuliano, A.J., Monsch, A.U. & Snow, M.G. (1995). Folstein vs modified mini-mental state examination in geriatric stroke: Stability, validity, and screening utility. *Archives of Neurology*, 52, 477-484.
- Graves, R.E., Bezeau, S.C., Fogarty, J. & Blair, R. (2004). Boston Naming Test short forms: A comparison of previous forms with new item response theory based forms. *Journal of Clinical and Experimental Neuropsychology*, 26, 891-902.

Grodzinsky, Y., Pin˜ango, M.M., Zurif, E & Drai, D. (1999). The critical role group studies in neuropsychology: Comprehension regularities in Broca's aphasia. *Brain and Language*, 67, 134 – 147.

Guedalia, J., Finkelstein, Y., Druker, A. & Frishberg, Y. (2000). The use of Luria's method for the neurobehavioral assessment of encephalopathy in an adolescent: application in a rehabilitation setting. *Archives of Clinical Neuropsychology*, 15, 177-184.

Heaton, R.K., Avitable, N., Grant, I., Matthews, C.G. (1999). Further crossvalidation of regression-based neuropsychological norms with an update for the Boston Naming Test. *Journal of Clinical and Experimental Neuropsychology*, 21, 572-582.

Heart and Stroke Foundation South Africa. (2007). Heart disease in South Africa. Retrieved November 25, 2007, from <http://www.heartfoundation.co.za/docs/heartmonth/HeartDiseaseinSA.pdf>

Hebben, N. & Milberg, W. (2002). *Essentials of neuropsychological assessment*. John New York: Wiley & Sons, Inc.

- Henry, G.K. (2001) The Rey Figure in Amazonia: Effects of jungle living in childrens' copy performance. *Developmental Neuropsychology*, 19, 33-39.
- Kalbe, E., Reinhold, N., Brand, M., Markowitsch, H.J. & Kessler, J. (2005). A new test battery to assess aphasic disturbances and associated cognitive dysfunctions – German normative data on the Aphasia Check List. *Clinical and Experimental Neuropsychology*, 27, 779-794.
- Kennepohl, S., Shore, D., Nabors, N., & Hanks, R. (2004). African American acculturation and neuropsychological test performance following traumatic brain injury. *Journal of the International Neuropsychological Society*, 10, 566-577.
- Killgore, W.D. & Adams, R.L. (1999). Prediction of Boston Naming Test performance from vocabulary scores: preliminary guidelines for interpretation. *Perception Motor Skills*, 89, 327 – 337.
- Kim, H. & Na, C.L. (1999). Normative data on the Korean version of the Boston Naming Test. *Journal of Clinical and Experimental Neuropsychology*, 21, 127-133.
- Kohnert, K.J., Hernandez, A.E., & Bates, E. (1998). Bilingual performance on the Boston Naming Test: Preliminary norms in Spanish and English. *Brain and Language*, 65, 422 – 440.

- Korkman, M. (1999). Applying Luria's diagnostic principles in the neuropsychological assessment of children. *Neuropsychology Review*, 9, 89-105.
- Kotik-Friedgut, B. (2006). Development of the Lurian approach: A cultural neurolinguistic perspective. *Neuropsychological Review*, 16, 43-52.
- Lachenicht, L. (2002). Correlation. In C. Tredoux & K. Durrheim (Eds.). *Numbers, hypotheses and conclusions A course in statistics for the social sciences*. (pp. 181 – 200)
- Lansing, A.E., Ivnik, R.J., Cullum, C.M., & Randolph, C. (1999). An empirically derived short form of the Boston Naming Test. *Archives of Clinical Neuropsychology*, 14, 481 – 487.
- Lezak, M.D. (1995). *Neuropsychological assessment* (3rd Ed.). New York: Oxford University Press.
- Luria, A.R. & Majovski, L.V. (1977). Basic approaches used in American and Soviet clinical neuropsychology. *American Psychologist*, 32, 959-968.
- Manly, J.J., Byrd, D.A., Touradji, P. & Stern, Y. (2004). Acculturation, reading level, and neuropsychological test performance among African American elders. *Applied Neuropsychology*, 11, 37 – 46.

- Manly, J.J. & Echemendia, R.J. (2007). Race-specific norms: Using the model of hypertension to understand issues of race, culture, and education in neuropsychology. *Archives of Clinical Neuropsychology*, 22, 319-325.
- Mansur, L.L., Radanovic, M., Taquemori, L., Greco, L. & Araujo, G.C. (2005). A study of the abilities in oral language comprehension of the Boston Diagnostic Aphasia Examination – Portuguese version: a reference guide for the Brazilian population. *Brazilian Journal of Medical and Biological Research*, 38, 277-292.
- Mba, C.J. (2004). The impact of external causes of South Africa's expectation of life. *African Population Studies*, 19, 165 – 177. Abstract retrieved September 15, 2005, from Bioline International.
- Mochan, A., Modi, M., & Modi, G. (2003). Stroke in black South African HIV-positive patients: A prospective analysis. *Stroke*, 34, 10-15.
- Mosdell, J. (2005). *Adaptations of aphasia tests for a South African context*. Unpublished honours thesis, University of Cape Town, South Africa.
- National Injury Mortality Surveillance System Annual Report. (2003). A profile of fatal injuries in South Africa. Retrieved April 7, 2005, from www.sahealthinfo.org/violence/nimss.htm

Nell, V. (1999). Luria in Uzbekistan: The vicissitudes of cross-cultural neuropsychology. *Neuropsychology Review*, 9, 47-52.

Nelson Mandela Study. (2002). South Africa HIV/AIDS statistics. Retrieved April 7, 2005, from <http://www.avert.org/safricastats.htm>

Nocentini, U., Goulet, P., Roberts, P.M. & Joannette, Y. (2001). The effects of left-versus right-hemisphere lesions in the sensitivity to intra- and interconceptual semantic relationships. *Neuropsychologia*, 39, 443-451.

Nysa, G.M.S., van Zandvoort, M.J.E., de Kort, P.L.M., Jansen, B.P.W., Kappelle, L.J. & de Haana, E.H.F. (2005). Restrictions of the Mini-Mental State Examination in acute stroke. *Archives of Clinical Neuropsychology*, 20, 623-629.

Odero, W., Garner, P. & Zwi, A. (1997). Road traffic injuries in developing countries: a comprehensive review of epidemiological studies. *Tropical Medicine & International Health*, 2, 445 – 460. Abstract received September 15, 2005, from Ingenta Connect.

Ostrosky, F., Ardila, A., Rosselli, M. Lopez-Arango, G., & Uriel-Mendoza, V. (1998). Neuropsychological test performance in illiterates. *Archives of Clinical Neuropsychology*, 13, 645 – 660.

- Ostrosky-Solís, F., López-Arango, G., & Ardila, A. (2000). Sensitivity and specificity of the Mini-Mental State Examination in a Spanish-speaking population. *Applied Neuropsychology, 7*, 25-31.
- Randolph, C., Lansing, A.E., Ivnik, R.J., Cullum, C.M. & Hermann, B.P. (1999). Determinants of confrontation naming performance. *Archives of Clinical Neuropsychology, 14*, 489 – 496.
- 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. *Aging and Mental Health, 7*, 217-223.
- Rapport, L.J., Millis, S.R. & Bonello, P.J. (1998). Validation of the Warrington theory of visual processing and the visual object and space perception battery. *Journal of Clinical and Experimental Neuropsychology, 20*, 211-220.
- Rey, G.J., Feldman, E., Rivas-Vazquez, R., Levin, B.E. & Benton, A. (1999). Neuropsychological test development and normative data on Hispanics. *Archives of Clinical Neuropsychology, 14*, 593-601.
- Rosselli, M. & Ardila, A. (2003). The impact of culture and education on non-verbal neuropsychological measurements: A critical review. *Brain and Cognition, 52*, 326-333.

Russell, E.W., Russell, S.L.K. & Hill, B.D. (2005). The fundamental psychometric status of neuropsychological batteries. *Archives of Clinical Neuropsychology*, 20, 785-794.

Spreeen, O. & Strauss, E. (1991). *A compendium of neuropsychological tests: Administration, norms, and commentary*. New York: Oxford University Press.

Stewart, R., Johnson, J., Richards, M., Brayne, C. & Mann, A. (2002). The distribution of Mini-Mental State Examination scores in an older UK African-Caribbean population compared to MRCCFA study norms. *International Journal of Geriatric Psychiatry*, 17, 745-751.

Storms, G., Saerens, J., & De Deyn, P.P. (2004). Normative data for the Boston Naming Test in native Dutch-speaking Belgian children and the relation with intelligence. *Brain and Language*, 91, 274 – 281.

Swartz, L. (1998). Language Diversity and mental health care. In L. Swartz (Ed.), *Culture and mental health: A South African view*. (pp. 25-51). Cape Town: Oxford University Press.

Swartz, L., Drennan, G., & Crawford, A. (1997). Changing Language policy in mental health services – a matter of interpretation? In D. Foster, M. Freeman and Y.

Pillay (Eds.), *Mental health policy issues for South Africa*. (pp.167-180). S.A.: Multimedia Publications.

Talberg, I.M. (2005). The Boston Naming Test in Swedish: Normative data. *Brain and Language*, 94, 19 – 31.

Trojano, L., Fragassi, N.A., Chiacchio, L., Izzo, O., Izzo, G., Di Cesere, G., Cristinzio, C. & Grossi, D. (2004). Relationships between constructional and visuospatial abilities in normal subjects and in focal brain-damaged patients. *Journal of Clinical and Experimental Neuropsychology*, 26, 1103-1112.

Tupper, D.E. (1999). Introduction: Alexander Luria's continuing influence on worldwide neuropsychology. *Neuropsychology Review*, 9, 1-7.

Walsh, K.W. (1991). *Neuropsychology A clinical approach*. London: Churchill Livingstone.

Walsh, K.W. (1991). *Understanding brain damage: A primer of neuropsychological evaluation*. London: Churchill Livingstone.

Wind, A.W., Schellevis, F.O.G., Van Staveren, G., Scholten, R.J.P.M., Jonker, C. & Van Eijk, J.T.M. (1997). Limitations of the Mini-Mental State Examination in diagnosing dementia in general practice. *International Journal of Geriatric Psychiatry*, 12, 101-108.

Wong, T.M. & Fujii, D.E. (2004). Neuropsychological assessment of Asian Americans: Demographic factors, cultural diversity, and practical guidelines. *Applied Neuropsychology, 11*, 23 – 36.

Appendices:

Appendix One

Screening for Controls

Please indicate whether you have had any of the following (either currently or previously):

Please be assured of the confidentiality and anonymity of any personal information that you give when participating in this study.

	<i>Please tick all that apply:</i>
A stroke	
A heart operation to treat coronary artery disease, e.g. coronary artery bypass graft surgery or stenting.	
Epilepsy or other seizures/fits	
A severe head injury	
Brain tumour or cancer	
Hydrocephalus ('water on the brain')	
Herpes encephalitis	
TB (tuberculosis)	
Diabetes (sugar disease)	
High blood pressure (hypertension)	
Multiple sclerosis	
Meningitis	
Alzheimer's disease	
Parkinson's disease	
Systemic Lupus Erythematosus (SLE)	
Syphilis	
HIV/AIDS	
Neurocysticercosis	
Recreational drugs (e.g. dagga, tik, cocaine etc)	
Any psychiatric condition (e.g. schizophrenia, bipolar)?	
Any other disease that you've had/currently have that may affect the brain?	

CONSENT FORM

TITLE OF PROJECT: A South African Neurocognitive Assessment Battery

Please cross out

as necessary

Have you read the Subject Information Sheet? YES/NO

Have you had an opportunity to ask questions and discuss the study? YES/NO

Have you received satisfactory answers to all your questions? YES/NO

Have you received enough information about the study? YES/NO

Who have you spoken to? Dr/Mr/Mrs/Ms/Prof.

Do you understand that you are free to withdraw from the study:

- at any time
- without having to give a reason for withdrawing
- and without affecting your future treatment? YES/NO

Do you consent to the unattributed and confidential use of these recordings for scientific purposes? YES/NO

Signed **Date:**

(NAME IN BLOCK LETTERS)

Appendix Three

Patient Information Sheet

TITLE OF PROJECT:

A South African Neurocognitive Assessment Battery

- You are invited to participate in a neuropsychological study conducted at Groote Schuur Hospital. Please read this information sheet carefully and do not hesitate to ask the researcher for any additional information.
- The overall purpose of the investigation is to adapt and validate a South African neurocognitive assessment battery, which comprises neuropsychological tests.
- You are asked to take part in this study by participating with different neuropsychological tests and tasks. You will be asked to attend two half-hour testing sessions a week apart.
- There are no anticipated risks involved in this research, but if you should experience mental and/or physical fatigue, or any form of psychological distress please be aware that you could inform the researcher immediately.
- It is up to you to decide whether or not you take part. If you decide to take part you will be given this information sheet to keep and asked to sign a consent form. If you decide to take part you are still free to withdraw from the study at any time, without having to give a reason and without this affecting future treatment.
- The confidentiality of your answers and your identity will be protected. All data collected will be suitably anonymous, securely stored, made accessible only to the researcher, and destroyed at the end of the project.
- This study is an educational project, forming part of a Ph.D. degree at the University of Cape Town (UCT). The research will be carried out by researchers from UCT and will be funded by the same university.
- The study has been reviewed by the UCT Psychology Department's ethics committee.
- If you have any questions regarding this study, or concerns regarding the manner in which the study was conducted, or would like to be informed of the results when the study is completed, please feel free to contact the principal researcher.
- Address for communications:

Professor Mark Solms:
Department of Psychology
University of Cape Town
Rondebosch 7701

Ph. (021) 650-3435

Appendix Four

Scoring sheet

Patient's name: _____ no. _____

1. Orientation

1.1 Person	
	<i>Normal score = 2/2</i>
1.2 Place	
	<i>Normal = 3/4</i>
1.3 Time	
	<i>Normal = 3/5</i>

2. Memory

2.1 Auditory Span

- Record patient's response on the lines provided:

2, 7 _____
5, 7, 2 _____
1, 9, 6, 4 _____
1, 4, 2, 7, 9 _____
8, 3, 7, 4, 6, 2 _____
7, 3, 5, 2, 4, 1, 9 _____

7 Digits = 2/2
5 Digits = 0/2
Normal = 6 digits

2.2 Four Hidden Objects

1. Show the 4 objects (a key, pipe, flower and bangle) and ask the patient to name them – *record items left out*,

4 Objects = 2/2
3 Objects = 1/2
2 Objects = 0/2
Normal = 4 objects

2. Hide all 4 objects in one location e.g. under the sheet and immediately ask the patient what the objects were – *record items left out*,

3. Distract the patient (eg. ask them about names and ages of children);

4. Repeat the question (NB: do not tell the patient how many objects or where you hid them);

4 Objects = 2/2
3 Objects = 1/2
2 Objects = 0/2
Normal = 4 objects

5. If patient fails, repeat the process. If patient successful, proceed to step 7 – *record how many repetitions*;

2.4 Rey Complex Figure - Copy

Name: _____

Date and time administered: _____

Rey Complex Copy – Second recall - immediate

Name: _____ Date and time administered: _____

Rey Complex Copy – Delayed recall at thirty minutes

Name: _____ Date and time administered: _____

Reading tests

Show me three fingers

.../1

Procedure: "I'm going to give you a short story. Please read it aloud."

Mary Selo story

Fluency: Make a note of time taken to read aloud story

Were all words properly enunciated? Make notes of problematic words

Other comments

.../2

Mild defect = 1

Normal = 2

Repetition

"Why am I sitting here?" (score =1)

"The painter painted many beautiful scenes " (cumulative score = 2)

"This doctor does not visit all the patients in the ward" (cumulative score =3)

"Why do the members of the committee not ask their representatives for aid?" (cumulative score=4)

Procedure: Compare to production

.../4

Normal = 3

Naming

Procedure: Ask patient to name body parts and objects at the bedside:

"Elbow, ankle, wrist, knee, shoulder"

.../5

Normal =4

"Pillow, sheet, spectacles, collar, buckle

.../5

Normal =4

Total

Naming Test:

Drawing shown:

Response from patient:

Incorrect/correct

3. Visual-spatial assessment

a. Neglect

Visual:

Tactile:

Auditory:

- (a) All stimuli correctly identified consistently in all modalities not affected by primary sensory impairment = 2/2
- (b) Consistently identifies both unilateral (L) and (R) stimuli correctly but frequently neglects (L) on bilateral stimulation (in modalities without primary sensory impairment) = 1/2
- (c) Consistently neglects (L) even on unilateral stimulation (in modalities without primary sensory impairment) = 0/2

4. Visual-spatial assessment

a. Anosognosia

If the patient does not spontaneously describe deficit, ask "Please describe all your current symptoms/deficits"

Score 3/3 if they can

If they do not describe deficit, ask "What about your legs/arms/hands/eyes, etc. (where applicable), are they all functioning normally?"

Score 2/3 if they can

If still denies deficit, demonstrate deficit to patient by physical examination, then ask: "Do you still think that your... is functioning normally?"

Score 1/3

Is there still a denial of deficit?

Score 0/3

5. Visual-spatial assessment

a. Hut drawing test

<i>Picture proportionate = 1</i> <i>Any sign that the left side of the picture is neglected = 0</i>
--

Hut drawing - copy

4. Executive

4.1 Controlled Oral Word Association Test

Procedure:

Ask patient to tell you as many words beginning with the letter F, A and S (or equivalent letters in Xhosa/Afrik). They must exclude proper names like the names of their friends and family or products like "Nike", must only be objects, give some examples.

Make a line for each 15 seconds to mark where the patient is. Allow 60 seconds for each trial.

->Please record repetitions with an "r" and rule breaks

F

A

S

Total F:

Total A:

Total S:

Total: _____

(a) More than 25 words (normal) = 2/2

(b) Between 15 and 25 words = 1/2

(c) Less than 15 words = 0/2

4.2 Red/Green

Perfect with one or two mistakes = 2

Consistently imperfect performance = 1

Gross-evidence of imperfect impulsivity, rule breaking, stereotyped responses = 0

4.3 Fist-Side-Palm

Perfect performance after one or initial errors = 1

Inability to achieve perfect performance despite repeated trials = 0

4.4 Tapping Test

Perfect performance after initial errors = 1

Inability to achieve criterion despite repeated trials = 0

4.5 Repeated Pattern Drawing



4.6 18 Book Problem

Correct response at first attempt = 2

Correct response after initial impulsive or stereotyped response = 1

Two incorrect responses = 0

Summary of Scores

	Subtotal	Total	Normal
Assessment of Orientation			
Orientation to Person			
Orientation to Place			
Orientation to Time		.../11	8
Assessment of Memory Function			
Digit Span			
Registration			
Simple Recall			
Complex Recall		.../18	14
Assessment of Language Function			
Production			
Comprehension			
Repetition			
Naming		.../30	26
Assessment of Spatial Cognition			
Spatial Cognition and Perception			
Neglect			
Anosognosia		.../15	12
Assessment of Frontal Function			
Deep White Matter			
Mesial			
Orbital / Basal			
Dorsolateral		.../11	10
Total		.../85	70

Appendix Five

Appendix Six

Appendix Seven

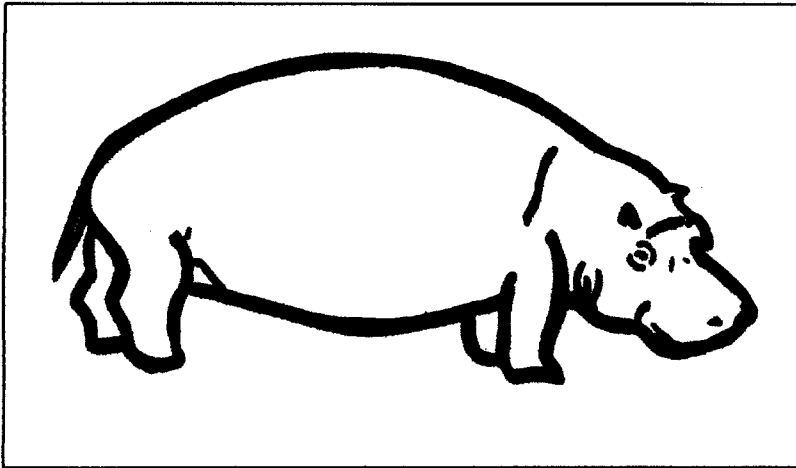


Figure 5. Hippo – problematic item in Naming Test.

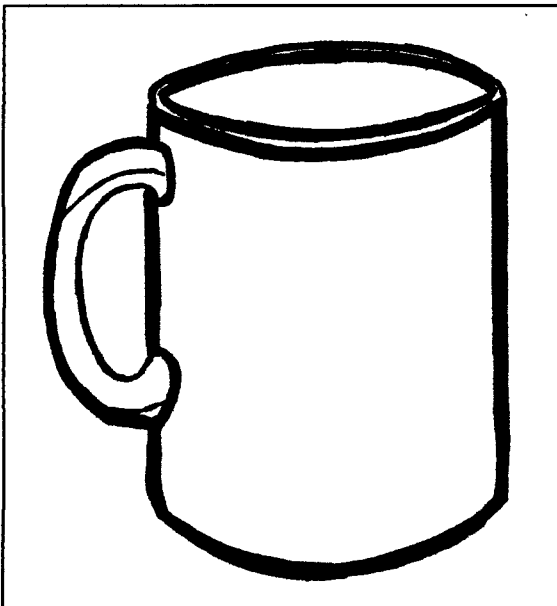
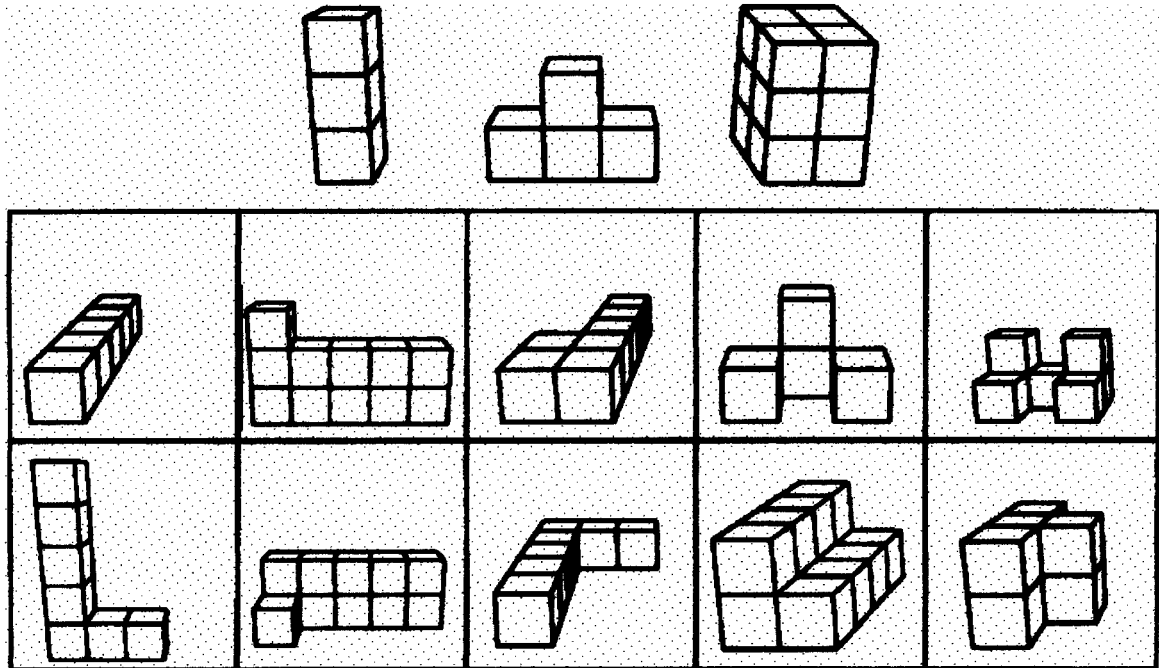


Figure 6. Mug – problematic item in Naming Test

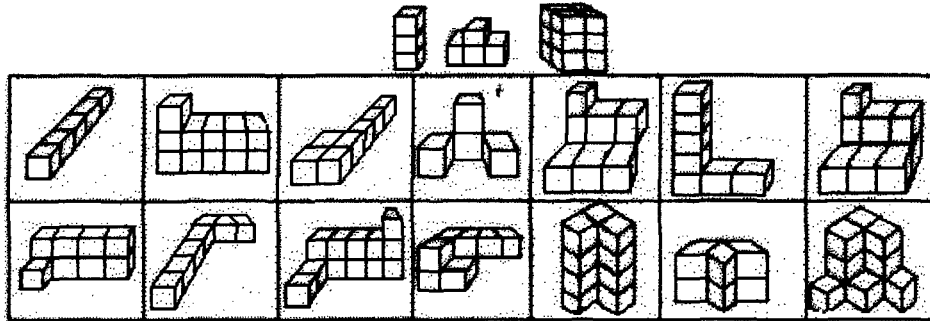
Appendix Eight

3-D Analysis Test



Appendix Nine

Old 3-D Analysis Test



Appendix Ten

