

THE CONSTRUCT VALIDITY AND RELIABILITY OF THE CHILD MEMORY SCALE, THE SEARCH AND MEMORY TASK, AND THE TOULOUSE-PIERON TEST FOR A SAMPLE OF SOUTH AFRICAN PRIMARY SCHOOL LEARNERS

**Hester Maria van der Merwe
401128**

A Research Report submitted to the Faculty of Humanities,
University of the Witwatersrand, Johannesburg, in partial fulfillment of
the requirements for the degree of Master of Education in Educational
Psychology by Coursework and Research Report.

Johannesburg, November 2009

Declaration

I hereby declare that this research report is my own independent work, and has not been presented for any other degree at any other academic institution, or published in any form.

It is submitted in partial fulfillment of the requirements for the degree of Master of Education in Educational Psychology by Coursework and Research Report at the University of the Witwatersrand, Johannesburg.

Rize van der Merwe

401128

30 November, 2009

For Johan

Acknowledgements

I wish to extend sincere thanks and grateful acknowledgement to the following:

Joseph Seabi, for his academic guidance.

Mike Greyling for conducting the computer based analysis, his invaluable advice on statistical theory, and his patience.

Dr Paul Goldschagg and the RANCH-SA team for including me in their study and allowing me to use the data for mine. Nicky and all the fourth year JCE students for their time and effort to get the test data captured in time for this study to take place. To Spiky Squares!

To all my family and friends for their financial and emotional support. Words are simply inept... A special word of thanks to my mother, Cecile; my remarkable sister Debbie & brother-in-law Louis; my brother Callie; Agribest Trust, and Theunis & Marietjie Prinsloo.

To Leon, who has dealt with my 'travels' far better than anyone I know.

Abstract

In an effort to contribute to growing knowledge about measuring instruments' applicability in different South African groups, the study attempted to ascertain the construct validity and reliability of three measuring instruments of memory. Furthermore, statistical properties and potential gender differences in a population of primary school learners in a school in KwaZulu-Natal between these instruments, namely the Child Memory Scale (CMS), the Search and Memory Task (SMT), and the Toulouse-Pieron Test (TPT), were explored. The internal reliability of the CMS was acceptable, but the Recognition section of the test was, on the whole, inconsistent and as such problematic. The SMT's reliability estimate was lower, but acceptable. Some of the items on the test proved to be problematic in terms of the test's overall reliability. Significant but weak relationships were found between the CMS and TPT, as well as between the SMT and TPT. However, no significant relationship was found between the SMT and the CMS. Furthermore, no significant gender differences were found between the three tests and no significant difference in correlation between the gender matrices was evident. These findings call the construct validity of the CMS, SMT and TPT in question, as literature (Gathercole & Martin, 1996; Logie, 1999) indicate probable significant inter-correlations between different aspects of memory.

Keywords: reliability, validity, Child Memory Scale, Search and Memory Task, Toulouse-Pieron Test.

Table of Contents

Acknowledgements	4
Abstract	5
CHAPTER ONE: INTRODUCTION	9
CHAPTER TWO: LITERATURE REVIEW	14
2.1 Brief Historic Overview	14
2.2 Test Bias	16
2.3 The role of culture, language and gender in psychological testing	20
2.4 Remedies for test bias	22
2.5 Validity and Reliability	23
2.6 Memory	24
2.6.1 Atkinson and Shiffrin Model	25
2.6.2 Baddeley and Hitch's Model of Memory	27
2.7 Gender differences	34
CHAPTER THREE: METHODS	36
3.1 Introduction	36
3.2 Research Design	36
3.3 Participants	36
3.4 Instruments	38
3.4.1 Demographic Questionnaire	38
3.4.2 Child Memory Scale	38
3.4.3 Search and Memory Task	39
3.4.4 Toulouse-Pieron Test	39
3.5 Procedure	40
3.6 Threats to Validity	42
3.7 Data Analysis	44
CHAPTER FOUR: RESULTS	45
CHAPTER FIVE: DISCUSSION	59
REFERENCE LIST	71
APPENDICES	
APPENDIX A: Parental Consent Form	
APPENDIX B: Participant Assent Form	
APPENDIX C: Child Questionnaire	
APPENDIX D: Child Memory Scale	
APPENDIX E: Search and Memory Test	
APPENDIX F: Toulouse-Pieron Test	
APPENDIX G: Ethical Clearance Certificate for the Study	

List of Tables

Table 1: Demographic Information	37
Table 2: Reliability Estimates for the Child Memory Scale and the Search and Memory Task	46
Table 3: Item analysis of the Child Memory Scale - Conceptual Recall	47
Table 4: Item analysis of the Child Memory Scale – Recognition Memory	48
Table 5: Item analysis of the Search and Memory Task	49
Table 6: Basic Descriptive Statistics for the Tests	52
Table 7: Kolmogorov-Smirnov Tests of Normality for the Tests	53
Table 8: Relationship between Child Memory Scale and the Search and Memory Task by Gender	54
Table 9: Relationship between the Child Memory Scale and the Toulouse-Pieron Test by Gender	55
Table 10: Relationship between the Search and Memory Task and the Toulouse-Pieron Test by Gender	56
Table 11: T-test for Independent Samples: Gender	57
Table 12: Fisher-Z Transformations on Gender correlations between the Child Memory Scale, Search and Memory Task, and Toulouse-Pieron Test	58

List of Figures

Figure 1: Histogram: CMS	50
Figure 2: Histogram: SMT	51
Figure 3: Histogram: TPT	51

CHAPTER ONE

INTRODUCTION

1.1 Introduction

The assessment of human intelligence, or parts thereof, has fascinated scientists for generations. To take an abstract notion, formulate it into an entity and then make it concrete by, for instance, using numbers to describe it seems to be a process that is an undeniable part of human kind (Gould, 1987). It appears inevitable that such a mindset would eventually lead to psychological testing. A psychological test is a standardized procedure for sampling behaviour, describing it with categories or scores according to norms and standards used to predict more significant behaviours (Gregory, 2007). The Standards for Educational and Psychological Tests of the American Educational Research Association, defines assessment as “a comprehensive evaluation of an individual through a process that integrates test information with information from other sources (e.g. from the individual’s social, educational, employment, or psychological history)” (AERA, 1999, p.3). A test thus measures specific construct(s) and may include inventories, scales and questionnaires, while assessment implies a more comprehensive undertaking, which may or may not include tests.

There is little doubt about the importance of accurate psychological assessment in the modern practice of psychology. Precise assessment is a basic and crucial step in effective evaluation, intervention, placement, selection, classification, promotion and research (Huysamen, 1980; Sehlapelo & Terre Blanche, 1996).

There are three areas of concern when considering psychometric tests: if a test truly measures what it claims to measure (validity), the consistency with which an instrument measures a construct (reliability), and if the results of such testing have been influenced by any other factors except for the construct measured (test bias) (Retief, 1988; Salkind, 2004). Reliability and validity have become almost standard for today's commonly accepted tests, but the existence of test bias is still a main concern (Foxcroft & Roodt, 2005).

Testing in South Africa presents with several challenges, especially taking into account its diverse society. As can be seen from the list of classified tests published by the Health Professions Council of South Africa (HPCSA), some original South African tests have been created, but the trend is still to standardize existing imported tests for the South African population (HPCSA, 2007). This inclination is understandable when considering the vast amount of resources with regards to time and money usually spent on the creation of an original test compared to standardizing an existing test. The HPCSA encourages practitioners to attend to the development and adaptation of culturally appropriate measures and also warns that appropriate research needs to be undertaken to investigate tests' cultural sensitivity. Particular care should also be taken when interpreting the results of such tests (HPCSA, 2007).

Currently, research is undertaken to study the effects of noise on learners' cognitive abilities. This study, the RANCH-SA, is a replication of a European cross-national study (RANCH-EU), and includes testing learners on memory, attention, and reading abilities (Stansfeld et al., 2005). The instruments used to determine these skills included the Child Memory Scale (CMS) which measures episodic memory and cued recall, the Search and Memory Task (SMT), which

measures working memory and attention, and the Toulouse-Pieron Test (TPT), which measures sustained visual attention. However, these measuring instruments have not yet been proven to be suitable for use in the South African context (HPCSA, 2007).

This study specifically attempts to examine the construct validity and reliability of the three instruments for both genders in the South African context as well as the statistical properties.

1.2 Rationale for the current study

According to the HPCSA listings (2007), there are no standardized memory tests for learners in South Africa at the moment. Most tests that measure memory and attention are intelligence tests and aptitude tests, for example the Wechsler Intelligence Scale for Children (WISC-III). In these cases, memory tasks are included as part of a test, not as the test *per se*. None of the abovementioned tests, i.e. CMS, SMT, and TPT, have been listed as standardised tests for the South African population, and no research has been conducted on their reliability for the South African context (HPCSA, 2007). Therefore, it is of particular interest to research the suitability of these tests for the South African population, and so to add to the existing body of standardised tests in South Africa.

The sample size used in the RANCH-SA study is too limited for a comprehensive standardization study. However, the collected data can be used to explore the internal consistency reliability of each of the three tests, as well as look for correlations between the

aspects of memory measured and compare correlations between the genders and thus examine the construct validity and psychometric properties of the instruments.

With the above literature review in mind, the study intended to examine the following:

Firstly, the reliability of the CMS, the SMT and the TPT for a sample of South African primary school learners.

Secondly, the construct validity of the CMS, the SMT and the TPT for a sample of South African school learners by: investigating the correlation between episodic memory and cued recall (as measured on the CMS), working memory and attention (as measured on the SMT), and sustained visual attention (as measured on the TPT); exploring the differences between working memory and attention; episodic memory and cued recall; and sustained visual attention between males and females; and examining a correlation between the correlation matrices of the male and female groups on working memory and attention, episodic memory and cued recall, and sustained visual attention.

The following hypotheses are put forward:

1. Significant inter-correlations will be found amongst episodic memory and sustained visual attention; working memory and attention, and sustained visual attention; episodic memory and cued recall, and working memory and attention.
2. There will be a significant difference between genders with regards to episodic memory and cued recall, in favour of females.

3. No statistically significant differences between males and females on working memory and attention will be found.
4. There will be a significant difference between genders with regards to sustained attention, in favour of males.
5. There will be a significant correlation between the correlation matrices of males and females.

CHAPTER TWO

LITERATURE REVIEW

2.1 Brief Historic Overview

In order to appreciate and even predict trends in modern psychological assessment, knowledge about its origins and contributing factors is crucial. As such, any discussion on psychometric tests and their influence in contemporary psychology would be incomplete without a critical look at the history of psychological assessment and practices, and specifically the role that South Africa's unique history and community plays in psychological assessment in this country.

The development of the South African psychological testing field closely followed international advances in psychological assessment, and especially the progress made in the United States of America, Britain and Western Europe since the 1800s (Huysamen, 1980). Long before any formal South African research was done, international researchers, such as De Gobineau in 1853 and Kingsley in 1899 already had conflicting views and interpretations of cross-cultural test results (Retief, 1988), which is of particular interest in the multi-cultural context of South Africa. Some attributed differences in test results to arrested development indicated by race (Fick, 1939), while others, such as Van den Berg (1938) argued that differences could be ascribed to cultural differences. It thus appears that there are two distinct directions regarding the issue of cross-cultural testing: those who attributed differences in test scores between cultural groups to genetic differences, implying that groups who were less successful on these tests were somehow unable

to thrive in western civilization, and those who attributed differences in tests scores to culture and life style differences (Klineberg, 1980).

As far back as 1915, researchers have been adapting measuring instruments to test different South African populations. Both Martin (1915) and Rich (1917) tested Zulu children with an adapted version of the Binet-Simon Scale, albeit with samples that differed slightly in age.

Biesheuvel (1949) insisted that, in order to test abilities in a specific culture, one should have a sound comprehension of the skills and knowledge of the population for which the test is intended, to ensure culturally appropriate items in a test. He also pointed out that traditional cultures were in transition as a result of having to adapt to Western influences (Biesheuvel, 1959).

Fick (1929; 1939) on the other hand, attributed differences in test scores to genetics, using these differences in test scores to prove ‘scientifically’ that Africans were inherently inferior to Europeans, and neglecting to take into account that the tests used were designed for Western populations. Views such as these of Fick, were supported during the Apartheids regime and the record of scientific development and application of psychometric measuring devices, instruments, methods and techniques in South Africa has thus been tainted by the effects of racial prejudice.

According to South Africa’s Bill of Human Rights, every person has the right to equality and dignity, and has the right to have these rights protected. Furthermore, everyone has the right to

physical and psychological integrity, implying, amongst other aspects, fair psychological testing. In order to promote and ensure fairness with regard to psychological testing, the HPCSA has mandated the Psychometrics Committee of the Professional Board of Psychology to classify tests and advise on their revision in order to be fair, reliable and valid in terms of what these psychological tests measure for specific populations. However, the legal requirements put the responsibility in the practitioner's court, ultimately making the test user accountable to ensure that tests used are valid for the populations for which they are intended. Currently the list of tests consists of 196 standardised tests, of which the majority is standardised imported tests (HPCSA, 2007). When taking into consideration the number of different cultural groups in South Africa, the list seems woefully inadequate.

2.2 Test Bias

Test bias is a controversial topic in psychological assessment, with some believing in the value and objectivity of psychological tests, and others rejecting the notions by proclaiming cultural and sexual bias against different groups, like ethnic minorities, women, and the poor. Much of the disagreement surrounding test bias is as a result of the interchangeable use in literature of two terms that refer to different aspects, namely 'fairness' and 'bias'. A distinction between 'fairness' and 'bias' is consequently in order.

Test fairness involves social values (Gregory, 2007). The perception of a test as fair, or not, is subjective and differs according to implicit and emotional assumptions of people that cause them to view information in diverse ways (Cole & Moss, 1989). According to the Standards for

Educational and Psychological Tests, there are four types of ways that the word ‘fairness’ is utilized in testing literature: fairness regarding selection on the grounds of a test; fairness in terms of outcomes of testing for different groups; fairness concerning treatment during the testing process; and fairness with reference to the opportunity to learn the material contained in the test (AERA, 1999).

Test bias, on the other hand, is a psychometric term that is quantifiable in that it can be measured mathematically and proven statistically (Retief, 1988). It is a technical concept that is objective and open to impartial scrutiny (Gregory, 2007). Retief (1988, p.44) defines bias as “...influences exercised by causes not related to the construct being assessed”. Consequently, an alternative source (not the construct being measured) is impacting on the test results and, by implication, members of different cultures or groups do not have an equal chance to answer a specific item correctly (Retief, 1988).

There is no established method to find possible sources of test bias. Retief (1988) maintains that it is a theoretical and inferential exercise where the researcher has to work his/her way back to possible causes. A sound and comprehensive knowledge of the culture in which testing is taking place seems to be a fundamental aspect of pinpointing possible test bias (Foxcroft, 2002). Sources of test bias can potentially be almost anything or combinations of any number of factors (Retief, 1988). Common sources include: item format, item content, test language, testing medium, test administration procedure, test instructions, test materials, time of testing, sampling procedures, test anxiety, test motivation, test fatigue, subject mood, response style, tester attributes, test environment, choice of test, and physical problems (Foxcroft, 2002; Retief, 1988).

Van der Vijver and Poortinga (1997) organize different sources of bias into three categories: construct bias, method bias, and item bias.

Construct bias refers to a difference in the definition of a psychological construct that is claimed to be measured in different groups. These concepts may be defined in diverse ways in African, Eastern and Western societies. For instance, intelligence in the Western sense of the word may refer to 'scholastic' intelligence, whereas in the African sense social intelligence may be emphasized to a larger extent (Greenfield, 1997). Should such a partial overlap between definitions of a specific construct in different cultural groups exist, there is the possibility of construct bias. It would therefore be prudent to clearly define the theoretical construct measured to specify its precise meaning. Differential appropriateness of behaviours associated with a construct within different cultures can also lead to construct bias. For instance, the behaviours attributed to being a good son or daughter in different cultural groups may be similar in some ways, but may differ vastly in other respects. Insufficient sampling of all aspects related to a construct, including instruments that are too short or that do not comprehensively cover all linked facets of said construct could also cause construct bias. (Van der Vijver & Poortinga, 1997; Van der Vijver & Tanzer, 2004)

Method bias refers to prejudice arising from the samples, instruments and administration of instruments. If samples are incomparable on characteristics other than the construct being measured, it may lead to biased results, for instance socio-demographic background factors such as formal educational background, socio-economic circumstances, cultural and family background, novelty of testing situation, recruitment procedures, and motivation can profoundly

influence test results. Physical problems, such as positive HIV status or AIDS, malnutrition and other illness also need to be considered, and their effect on test results ascertained (Foxcroft & Roodt, 2005). The venue and time of testing seems to be of importance in terms of test bias, as well. To write a test in an unfamiliar environment that is extremely foreign to the test taker, may increase the test taker's anxiety levels so as to interfere with test results. Sensitivity needs to be shown with regards to the time of testing, since traditions and religious observances have to be taken into account (Foxcroft, 2002). Instrument bias refers to bias as a result of instrument traits, such as stimulus familiarity, differential familiarity with response procedures, differential response styles such as acquiescence and extremity ratings may influence test results unduly. Administration problems, such as differences in environmental administration conditions, physical and social, different levels of expertise of administrators, tester effects, communication problems, familiarity of stimulus material, differential familiarity with response procedures, and differential response styles, may also impact test results unduly. In addition, ambiguous instructions for respondents, unclear guidelines for administrators, the culture, appearance, obtrusiveness of the tester, and communication and/or language are all possible sources of administration bias (Foxcroft, 2002; Van der Vijver & Tanzer, 2004). The effect of the test administrator (Retief, 1988), lack of test-wiseness (Nell, 1999), and hesitance to request help during group testing as a result of cultural differences (Foxcroft, 2002), could further impact test results unduly. Method bias, defined as sample, instrument and administration bias, can comprehensively influence observable differences of test scores between cultural groups.

Item bias denotes complications at an item level. Poor item translation and ambiguous items, nuisance factors and cultural specifics, like incidental differences in connotative meanings and

inappropriate item content, are common sources of item bias (Van der Vijver & Poortinga, 1997). When a specific item in a test has different psychological connotations for different cultural groups, a comparison between groups on the construct measured by the item would prove to be worthless (Van der Vijver & Tanzer, 2004).

2.3 The role of culture, language and gender in psychological testing

Culture, language and gender play a role in almost every aspect of the testing situation.

Definitions for culture seem to focus on different aspects of culture and vary considerably.

Hofstede (n.d.) defines culture as the “collective programming of the human mind that distinguishes the members of one human group from those of another”.

Schein (1997) describes culture as shared, basic, unconsciously-held assumptions and beliefs of a group of people that assist them in defining themselves within their environment.

The Roshan Cultural Heritage Institute (2001) defines it as ‘Ways of Life’, including language, arts and sciences, spirituality, social activity and interaction.

Veroff and Goldberger (1995) specifically mentions that people in a cultural group often share a common history and live in a specific geographic area.

It appears that most definitions of culture seem to have the following in common:

1. It refers to the sum total of shared symbolic patterns, beliefs, values and attitudes that typically influences perception and behaviour in a specific group.
2. It is learned and mutually constructed through a perpetual process of social interaction that has been passed on through generations.
3. Although it is perceived as 'natural', it is in fact arbitrary, that is, not based on 'natural laws', but is instead created by members of a group.

Differences between cultural groups obviously need to be taken in consideration when planning, administering and interpreting psychological tests. When interpreting test results, appropriate norms have to be consulted. Unfortunately, many tests have not been standardised to South African populations, making it unethical to apply them to South African samples (HPCSA, 2007). Also, although the tracking of chronological age is important in Western societies, many other cultural groups do not consider it important. This could lead to complications when the exact age of a child is required to interpret a test (Foxcroft, 2002), for example, an intelligence test. Furthermore, it is possible for an unbiased test to be considered unfair on the grounds of social consequences of it being used for selection decisions (Gregory, 2007). As individual differences do exist within specific cultures, practitioners should also be aware of the pitfall of stereotypical overgeneralization. Ultimately, clinical test results need to be interpreted in the context of the individual (Gregory, 2007).

2.4 Remedies for test bias

In literature, bias is considered undesirable and most techniques focus on the reduction or elimination of it. However, it can also be viewed as an opportunity to learn more about cross-cultural differences (Van de Vijver & Tanzer, 2004). Possible remedies for construct bias are cultural decentering (identifying words and concepts that are specific to a particular language or culture and eliminating them), the convergence approach (instruments developed within a culture and subsequent cross-cultural administration of all instruments), the use of informants with extensive knowledge of the local culture and language; local surveys; non-standard instrument administration; cross-cultural comparison of nomological networks; and a sample of bilingual subjects, although the last techniques does appear to have limitations and a combination between bilingual and monolingual samples may be more useful (Van de Vijver & Tanzer, 2004). Method bias can be addressed by extensive training of the administrators; a detailed manual for administration, scoring and interpretation; comprehensive instructions and examples; use of subject and context variables, such as educational background; use of collateral information, such as test-taking behaviour or test-taking attitudes; assessment of response styles; and use of test-retest, training and/or intervention studies. Item bias can be attended to via judgmental methods of item bias detection, such as psychological analysis; psychometric methods of item bias detection; error or distracter analysis; and documentation of 'spare items' in the test manual that are equally effective measures of the measured construct as actually used items in the test (Van de Vijver & Tanzer, 2004).

2.5 Validity and Reliability

The question of bias goes hand in hand with the issue of validity. Traditionally, three types of validity are accepted by experts, namely content, criterion-related, and construct validity (Gregory, 2007). Content validity refers to the degree to which items in a test cover all the content areas pertaining to the construct being measured (Rosnow & Rosenthal, 2008). In other words, does the sample of items in the measuring instrument reflect all possible items on a specific topic (Salkind, 2004)? Criterion-related validity refers to the degree to which test results on a measuring instrument correlates with another outcome with which they (test results) should reasonably be related (Rosnow & Rosenthal, 2008). This study focuses on construct validity, which refers to whether or not a test measures some fundamental psychological construct in a measuring instrument or tool, and is thus based on an underlying construct (consisting of a group of interrelated variables) behind a measuring instrument or tool (Salkind, 2004). Construct validity indicates what the test truly measures, and as such an instrument's capacity to discriminate between different constructs is an important aspect of its construct validity (Rosnow & Rosenthal, 2008). Two other types of validation evidence were suggested by Campbell and Fiske (1959) as important in determining construct validation: convergent validity (in other words, correlation between the test results of two different measures of the same construct) and discriminant validity (distinctiveness of related but different measures). Construct validity can be determined by correlating a set of test scores with a theorized outcome that reflects the construct for which the test is designed (Salkind, 2004).

Reliability is defined as the degree to which test results can be generalized over testing situations, test forms and test populations (Huysamen, 1994). Different types of reliability include test-retest reliability, parallel forms reliability, internal consistency reliability, and inter-rater reliability. For the purposes of this report the focus is on internal consistency reliability, which refers to the degree of consistency items of the same test exhibit with regard to the construct that they measure (Salkind, 2004). This type of reliability can be determined by computing Cronbach's Alpha, split-half reliability, Spearman-Brown, Kuder-Richardson 20 and 21, to name but a few methods (Salkind, 2004).

Cognisance of threats to validity and test bias is crucial when considering the reliability of a measure for a specific context. In order to evaluate the reliability of a memory test to a certain population, a comprehensive overview of the construct being measured - memory - seems essential.

2.6 Memory

Defining memory is complicated by the fact that the word 'memory' is used for many different concepts. 'Memory' often refers to a neuro-cognitive capacity to encode, store, and retrieve information; a hypothetical store in which information is held; the information in the store; some property of that information; a componential process of retrieval of that information, and an individual's phenomenal awareness of remembering something. Other uses of 'memory' are largely idiosyncratic (Bower, 2000). Lezak (1983) defines memory simply as "...the means by which an organism registers some previous exposure to an event or experience" (Lezak, 1983 p.

414), and proceeds to explain that it includes several different functions like instantaneous retention; learning in terms of extent of recent memory, learning capability, retention of newly learned material; and efficiency of recovery of both newly learned and long-stored information (remote memory) (Lezak, 1983).

Literature on memory seems to be divided in two groups: those who believe in a unitary memory, and those who believe in multiple memories (Bower, 2000). The unitarian view is that memory is one, but that it can be measured in different ways, and that all of memory can be described. On the other hand, some believe that memory is comprised of multiple systems and attempt to describe and understand every type of memory system. According to this view, which currently seems to be the most widely accepted, memory is largely seen as a non-unitary entity, consistent of various memory systems. Schacter and Tulving (1994) suggested five main memory systems: working memory, episodic memory, semantic memory, perceptual representation system and procedural memory.

2.6.1 Atkinson and Shiffrin model

Most theories of memory consider either the structure of the memory system or memory processes or both. Structure refers in this case to the organization of the memory system whereas memory processes entail those actions that take place within the structure (Eysenck & Keane, 2000). Initially, many theorists focused on defining the structure of memory, such as the Atkinson and Shiffrin theory of memory (1968) which implies an inflexible storage structure consisting of sensory memory, short-term memory, and long-term memory. It is an almost linear

modal model where information received from the senses is very briefly held in the sensory store, divided in an iconic store and an echoic store, and if this information is attended to, it is then transferred to the short-term store, which has a very limited capacity. Mainly through repetition, information is transferred to the long-term store, which is defined as having an unlimited facility (Atkinson & Shiffrin, 1968). There are several problems with regards to this model. Firstly, it oversimplifies a complicated matter such as memory and considers both short- and long-term store as unitary concepts, a notion disproved by several researchers, of which Shallice and Warrington (1970) are the most prominent. The researchers tested and described the memory functioning of a patient with a grossly defective short-term store and found that although one would expect (according to the Atkinson & Shiffrin model) a general level of impairment that includes defective learning, memory and comprehension, no such result was found. This seriously calls the notion of the short-term store as a working memory, into question. The linear progression of a memory from the sensory to the short-term to the long-term memory, also poses a problem. Information that advances from the sensory memory to the short-term memory, have already been in contact with information in the long-term memory, as in the case of pronunciation of words where information from the long-term memory is needed in order for processing in the short-term memory to take place (Gathercole & Martin 1996; Logie, 1999). Perhaps this model focuses too much on the structure of memory as opposed to the processes involved in it.

2.6.2 Baddeley and Hitch's Model of Memory

Baddeley and Hitch (1974) formulated an alternative model for memory, specifically focusing on working memory. The idea of working memory was first suggested by Miller, Galanter and Pribram (1960). They defined working memory as the practical component of the short-term memory. It functions as part of an integrated system for holding and manipulating information during complex operations (Miller et al, 1960). Working memory has a limited capacity and is temporary in nature. The contents of working memory can dynamically be combined and manipulated with pre-existing information from the long-term memory to form new knowledge (Logie, 1999).

The Baddeley and Hitch model (1974) provided a broad theoretical framework that attempted to explain certain known phenomena but left ample space for further exploration and more specific sub-models. It defined memory as comprising of several components and focused on the function of the system in overall complex cognitions. According to the model, working memory consists of two slave systems, namely a phonemic or articulatory loop associated with verbal and auditory tasks, and a comparable system for visual memory that later was called the visuo-spatial scratch pad. A modality-free regulation system, called the central executive functions, control these two slave systems. These three components were thought to work independently of each other and have limited capabilities.

Initially the phonological loop, associated with the storage of acoustic and verbal information, was considered a very simple system (Baddeley & Hitch, 1974). However, Baddeley (1986;

1997) describes two components: a storage component or the phonological store that is able to hold verbal based information and that is accessible through auditory presentation or articulatory coding of visual material, and a rehearsal component or articulatory control process that is based on inner speech and thus is involved in subvocal rehearsal. Memory traces in the phonological store have a limited 'life span' of one and a half to two seconds, that is, unless it is refreshed by virtue of a process of subvocal rehearsal (Baddeley, 1997; 2003). The phonological loop has a limited capacity and as the number of items to be remembered increases, the first items start to fade before it can be rehearsed. Written material can be converted into a phonological code and registered in the phonological store by the articulatory control process. It has been established that the phonological loop plays an important role in learning to read, the comprehension of language and in acquiring vocabulary in young children (Gathercole & Baddeley, 1989; 1990a; 1990b; Gathercole & Martin, 1996; Jones, Gobet, Pine, 2005; Jones, Gobet, & Pine, 2007). Baddeley, Gathercole and Papagno (1998) go as far as to state that the phonological loop's primary role is to store unfamiliar sounds and words while they are being constructed in a more permanent configuration and attribute less importance to its role in sequences of familiar words. Language habits seem to have a profound effect on performance through the articulatory aspect of the phonological loop, while the storage component is relatively free from influence by language. Therefore, they postulate that language acquisition is facilitated by the phonological loop in two ways, namely by a store that can temporarily hold new sounds/words freely, and an articulatory control process that promotes learning through rehearsal. A dynamic interaction seems to be taking place between the loop and the long-term memory, as in the case of language acquisition, yet the precise mechanism that makes this possible is unclear (Baddeley, 2003). Baddeley and Hitch's version of the phonological loop does not, however, give an indication of

how serial order of incoming items is maintained (Baddeley, 2002). The parts of the brain that are implicated in the phonological loop, are the left temporoparietal region (Vallar & Papagno, 2002), especially BA 40 implicated in the storage component and Broca's area implicated in the rehearsal component (Awh, Jonides, Smith, Schumacher, Koeppel, Katz, 1996; Smith, Jonides, & Koeppel, 1996). According to Clark et al. (2001), the following areas in the brain are connected to verbal working memory: the occipital pole is linked with early visual processing; the superior temporal and inferior parietal gyrus bilaterally and the left angular gyrus is associated with visual and phonological word processing; the dorsal lateral occipital gyrus is implicated in spatial processing; and aspects of the bilateral superior parietal lobe are connected with imagery and episodic verbal memory. The lateral and superior prefrontal regions are linked to working memory control of sensori-motor processes.

The visuo-spatial sketchpad (formerly known as a 'scratchpad') acts as the phonological loop's visual and spatial equivalent, and is considered to temporarily hold and manipulate visuospatial information (Logie, 1995). As such, it plays a major role in resolution of visuospatial challenges and spatial orientation. It connects visual information with similar motor and haptic information and is accessed either through the senses or from the long-term memory (Baddeley, 2002). Logie (1995) suggests that the visual-spatial scratch-pad can be divided into a visual store component as well as a rehearsal aspect, and that the rehearsal mechanism for the visual spatial sketchpad is the inner scribe, which is the spatial component. Darling, Della Sala, Logie, and Cantagallo (2006) suggest that distinct cognitive systems underlie the visuo-spatial tasks and differences in type of information, appearance and location favoured by the systems may be at the heart of this separation. Baddeley (2002) believes that the central executive, through a

process similar to attention, plays a role in maintenance rehearsal in the visuospatial sketchpad. It is speculated that the visual-spatial sketchpad plays a role in acquiring semantic knowledge about objects, their appearance, how they can be manipulated for use, understanding complex systems such as machines, spatial orientation and geographical knowledge (Baddeley, 2003). According to PET evidence, the occipital parietal, prefrontal, and frontal lobes have been linked to visuo-spatial tasks (Jonides, Smith, Koeppe, Awh, Monoshima, & Mintun; 1993). Anatomically, the region of the brain indicated in visuo-spatial functioning appears to largely be the right hemisphere (Smith et al., 1996). The occipital cortex, the parietal cortex, and the premotor cortex have been indicated in spatial tasks.

The central executive plays a crucial role in the working memory system in that it controls and regulates the two slave systems, but does not have a storage capacity. It reflects on information from the environment as well as episodic memory to create a mental representation of a situation and then selects an appropriate response in reaction to it. The central executive is in part an attentional system that controls the focus, division and switch of attention (Baddeley, 1996), but also connects working memory to the long-term memory (Baddeley, 2002). This is done via a newer component to the Baddeley and Hitch model called the 'episodic buffer'. The episodic buffer was proposed by Baddeley (2000) as an addition in order to clarify certain previously inexplicable phenomena, such as the effect of articulatory suppression, and 'chunking' in the recall of prose. It is able to account for the integration and interaction of the phonological and visuo-spatial systems with each other, as well as with the long-term memory (Baddeley, 2003). Multidimensional coding makes this integration possible and its accessibility to conscious awareness offer a vehicle for binding and retrieval of information to ultimately form integrated

‘episodes’. The episodic buffer has a limited storage facility and is controlled by the central executive functions (Baddeley, 2003). The central executive is likely to incorporate many different areas of the brain in order to form a network rather than be limited to a specific area of the brain, as it involves both unity and diversity of processes (Garavan, Ross, Li & Stein, 2000).

Confirmation of a multi-component working memory system in children comes from several researchers (Alloway, Gathercole, Willis, & Adams, 2004; Gathercole & Hitch, 1993; Gathercole, Pickering, Ambridge, & Wearing, 2004). It appears that the three components of working memory as defined by Baddeley and Hitch (1974) are established in children by the age of six and that these functions increase linearly in terms of capacity from the age of four into early adolescence. According to Gathercole et al. (2004), the working memory structure seems to remain constant during childhood, however, their study makes no mention of the episodic buffer. Alloway et al. (2004) take it one step further and also state the existence of the episodic buffer in the working memory of children between the ages of four and six.

The Baddeley and Hitch model provides an account of the integrated and interactive system for momentarily storing and manipulating information in the human mind. It consists of several subsystems that work independently but in an integrated manner, mediated by the executive function and the episodic buffer, which appear to integrate information from the different memory systems. One of the limits of the model is that it fails to explain conative, emotional and motivational control of the system, in other words, mental energy/arousal (Baddeley, 2003). This is an important oversight, especially considering Bion’s theory of thinking, in which he claims that the earliest form of thinking, which forms the basis of later stages of thinking, is the

result of early emotional events between mother and infant and is decisive for the establishment of the infant's capacity to think (O'Shaughnessy, 2004). One would have to admit that any model not considering the emotional and motivational aspects, is incomplete. In this study working memory is measured by the SMT.

An aspect that plays a vital role in short-term memory is attention. Attention refers to "selectivity of processing" (Eysenck & Keane, 2000; p.119). The first theories that attempted to explain attention, were the filter theories, such as Broadbent (1958)'s theory, according to which a filter system selects whether to pay attention to incoming stimuli, organized in channels. Capacity theories followed, such as suggested by Kahneman (1973), according to which it is suggested that every person has a limited quantity of cognitive resources to expend on specific tasks. A major function of vision is object recognition and this is often achieved by limiting the elements that needs to be attended to and thus selecting a limited visual area in which to operate. This strategy is used repeatedly and the mechanism by which it is done, is called visual attention. Visual attention can be either goal-driven or stimulus-driven. The term 'goal-driven' implies action or intentional control on the side of the individual and he/she employs strategies to control attention. 'Stimulus-driven', on the other hand, suggests that attention is controlled by a feature of the situation in which the individual finds himself, that is, a prominent aspect of the object to be observed plays a major role and is not necessarily related to the individual's goals. One such example is a big, shiny limousine in the street – the eye is almost involuntarily drawn towards it. Often, the two types of control are combined in visual observation, even complementing each other, leading to comprehensive visual performance. Attentional control thus stems from interaction between the viewer's intent and the image's attributes (Yantis, 1998).

In the literature there are mainly two theories regarding visual attention: the spot light approach (Posner, 1980; Posner, Snyder, & Davidson, 1980) and the zoom lens approach (LaBerge, 1983). According to the spot light approach, attention is similar to a spot light and everything within a limited visual area is clearly visible, whereas anything falling outside of this area is much more difficult to see. Posner (1980) and Pashler (1998) argue that “covert attention” may exist, where the attention field can shift without any eye movement. According to the zoom lens approach, attention can be focused on a specific area in the visual field and can be increased or decreased according to the task at hand (LaBerge, 1983). Hoffman (1998) states that attention is vital in guiding the eye to pertinent areas of a scene/representation. However, focused visual attention appears to be more complex than either model imply. Visual attention is not always directed towards regions but can also be focused on objects (Marshall & Halligan, 1994; Neisser & Becklen, 1975). In this study visual attention will be measured by the TPT.

Episodic memory and semantic memory are memory systems that are part of long-term memory and have been described by some researchers in terms of the information they hold (Tulving, 1972) and by others in terms of the nature of subjective experience during encoding and retrieval (Wheeler, Stuss, & Tulving, 1997). Episodic memory was first defined as the storage and retrieval of specific events at specific times (Tulving, 1972), and now is seen as involving a special kind of awareness that can be described as “autonoetic” or self-knowing. This refers to the awareness experienced when thinking of a specific time in one’s personal history and consciously remembering a past episode as it was previously experienced, and is as such subjective (Wheeler et al, 1997).

Semantic memory has been defined as “ the mental thesaurus, organized knowledge a person possesses about words and other verbal symbols, their meaning and referents, about relations among them, and about rules, formulas, and algorithms for the manipulation of these symbols, concepts, and relations” (Tulving, 1972, p. 386). Wheeler et al. (1997) describe semantic memories as involving “noetic” or knowing awareness, in other words, a person thinks objectively about facts that he/she knows. Although there are major differences between these two types of memory systems, the way information is stored in the memories is exceedingly analogous. When information is encoded, invariably it is stored in both kinds of memory (Eysenck & Keane, 2000). Both episodic and semantic memory will be measured by the CMS in this study.

2.7 Gender differences

There seem to be gender differences with regard to different aspects of memory. Women generally outperform men on episodic memory tasks of verbal nature, in other words, when the items to be remembered are words (Hill, et al., 1995) or stories (Hultsch, Masson, & Small, 1991). However, with regard to visuospatial episodic memory tasks, men outperform women (Lewin, Wolgers, & Herlitz, 2001). Gender differences concerning working memory are complex. In a study by Loring-Meier and Halpern (1999), no gender differences regarding the accuracy of visuospatial working memory as measured on four tasks, were found. Men, however, tended to respond more rapidly on the tasks, leading the investigators to deduce that men may be more proficient in such tasks. The researchers drew the conclusion that tasks where speediness plays a role, may put females at a disadvantage. According to Giambra and Quilter

(1989), females are also at a disadvantage with regard to tests of sustained attention, and response time was once again slower than for males.

As already mentioned, memory is normally measured as part of an intelligence or aptitude test. Although several of these tests are available, a test that measures only memory or a function of memory has yet to be standardized for the South African context. The CMS, the SMT, and the TPT have been used in studies elsewhere, but have not hitherto been conducted in South Africa.

Stansfeld et al. (2005) used all three tests in cross-national research on the effects of aircraft and road traffic noise on learners' cognitions as part of the Road Traffic and Aircraft Noise Exposure and Children's Cognition and Health: Exposure-effect Relationships and Combined Effects (RANCH) project. The CMS, a normed and statistically valid long-term memory task, was used to measure delayed cued recall and recognition. The SMT was used to measure working memory and attention and sustained visual attention was measured using the TPT, a timed test based on simple visual figures. These tests have been used extensively in other studies (Hygge, Boman, & Enmarker, 2003; Gomes, Martinho, & Castelo, 1999; Lopez & Vazques, 2003; Matsui, Stansfeld, Haines & Head, 2004; Smith & Miles, 1987), and are considered to be reliable.

A study is being conducted with a view to replicate the abovementioned study, using the same protocol and tests. This research project, the RANCH-SA study, provides an opportunity to investigate the statistical properties, reliability and validity of these tests for South African populations.

CHAPTER THREE

METHOD

3.1 Introduction

This chapter describes the research design, sampling strategy, demographic particulars of the sample, the measuring instruments, namely the Child Demographic Questionnaire (CDQ), Child Memory Scale (CMS), Search and Memory Task (SMT), and Toulouse-Pieron Test (TPT); as well as the data collection protocol, threats of validity and data analysis.

3.2 Research Design

A cross-sectional design was used in the study.

3.3 Participants

A purposive sampling approach was used as certain prerequisites had to be met by participants in the study and as a consequence Grade Five and Six learners at a school were invited to take part in the research project as they already have a measure of ‘test-wiseness’, having been in school for at least four years. Test-wiseness, which can influence test results unduly, is acquired directly through repeated exposure to psychological tests and school examinations, and indirectly by acquiring the ‘skills’ necessary to function in a classroom, such as the capability to sit still and pay attention for long periods of time, the adroitness with which a child used pen and paper,

the ability to copy designs and solve problems, and work fast to keep up with classmates and to complete examination within the allotted time (Nell, 1999). A sample of 204 learners eventually participated.

The population of learners in the school appeared to be mostly of Zulu and Indian origin. As illustrated in Table One, sixty-one (32.6%) of them indicated English as the home language, while 63 (33.6%) and 59 (31.6%) of these learners indicated Zulu and English as the languages spoken at home, respectively. The female learners (n=102) outnumbered the male learners (n=85) to some extent. The learners fell in the age range of 10-14 years, although most ranged from 10-12 years in age, with one learner aged nine, three learners aged 13 and two learners aged 14. The mean age was 11.4 years. The language of instruction in school is English. Socio-economic status of learners was estimated by participation in the school's nutrition scheme, and a total of 74 learners (40%) were provided with a meal at school.

Table 1: Demographic Information

N	Males	Females	Mean Age <i>(in years)</i>	Age Range <i>(in years)</i>	English	Zulu	English/ Zulu	Other Languages
187	85	102	11.4	10-14	61	63	59	4

3.4 Instruments:

As part of a test battery, the CDQ, CMS, SMT, and TPT were administered to the learners. The four instruments are described below.

3.4.1 Child Demographic Questionnaire (CDQ)

The demographic questionnaire was designed to gain information on particular demographic variables, such as age, gender, and home language. The original questionnaire was adapted in terms of demographic detail, such as languages spoken at home, to better suit a South African population. It consisted of close-ended questions, and was completed by the participants.

3.4.2 The Child Memory Scale (CMS)

The CMS is used to assess episodic memory recognition and recall, thus the ability to comprehend, encode and recall meaningful and semantically related verbal material that is presented in a sequential format. The test consists of two stories, about a paragraph in length. These are read to the learners, and 30 minutes later they are given a cued recall task in which there are questions relating to the paragraphs. The test takes about 20 minutes to administer. An adapted edition of the original test, identical to the version used in the RANCH-EU study, was used. The adaptation included cued recall questions to make it suitable for group administration. The test was called ‘Story Memory’ and can be used to test learners from five to 16 years old.

The CMS has been used in several studies abroad (Cohen, 1997; Matsui et al., 2004) and has proved to be both reliable and valid for a sample of German learners (Matsui et al., 2004).

3.4.3 Search and Memory Task (SMT)

The SMT is used to test working memory and attention, and learners have to memorise between one and seven letters and are then asked to search through lines of jumbled letters for items that match the targets. The testing time is 10 minutes. An adapted version of the test that included questions of graduated intricacy, was used. The test is informally known as ‘The Letter Detective’. The test has been used in a number of international studies and appears to be both valid and reliable in the various samples used in the studies (Hygge et al., 2003; Smith & Miles, 1987).

3.4.4 Toulouse-Pieron Test (TPT)

The TPT demands great concentration and evaluates the aptitude or capacity to concentrate on tasks that are monotonous. Testees are expected to identify exemplars of two target figures which were randomly distributed within a large set of similar, but different, figures. There is a total of 1,600 figures of which 400 are targets. These figures are arranged in 40 rows that each consist of 40 figures and with ten target figures in each row. The test is informally known as ‘Spiky Squares’, has been used in a variety of studies and emerges as both valid and reliable for the specific samples used in these studies (Gomes et al., 1999; López & Vazques, 2003).

3.5 Procedure:

After receiving clearance from the Department of Education as well as the Committee for Research on Human Subjects- Humanities (Protocol: 2008ECE94), the researcher started off by planning and executing a pilot study. Information and experience gained by this study was used to enhance the main research project. The principals of the preselected schools were approached and their collaboration and written consent procured. Information regarding the classes and school schedule was obtained. The original tests, contact details of the participating schools and maps to the schools were prepared in advance and kept on file and on computer disk. Test administrators were trained in advance in the use of the tests and in standard points of protocol testing, such as the importance of keeping the test situation as much the same for the different classes as possible by following the same verbal instructions in each of the classes and not giving more positive encouragement to one class over another. The role of the administrator was made clear and included leader duties and general class monitor tasks. Other general guidelines directed conduct of the administrator during the test and specifically included a sensitive approach to learners. The test materials were prepared beforehand, identification codes including the school, class and pupil, were allocated and coded class lists generated. Informed, written consent from the learners involved in the study, and their parents was obtained (Appendices A & B). No learner was allowed to participate unless an assent letter had been completed and handed to the researcher.

Grade Five and Six learners from the school were tested with the exact same test battery as used in the RANCH-EU study. Attention was paid to making sure that each team of administrators

included people from different cultural backgrounds and genders, to minimize possible administrator bias (Foxcroft, Roodt, & Abrahams, 2005). The administrators introduced themselves and the study to the class, using the script from the in-school manual, carefully avoiding the word 'noise', so as not to unduly influencing the results. Child packs, envelopes containing all relevant test material, were distributed according to the class lists which gave the administrators a chance to take roll call and note the reason for absence of any child. Any latecomers were unable to complete the CMS, but completed the rest of the tests. The learners were also informed of their right to confidentiality and their right not to answer a question that they felt uncomfortable with answering. Any questions were fielded. The utmost care was taken to keep test responses confidential. To this end, the completed tests were put in an envelope immediately after completion and all collected information was coded and kept on a password protected computer. Only the researcher and the supervisor had access to this information.

Ahead of completion of each test, sample items were completed to ensure that participants understood what was expected of them and any questions regarding the work was answered. According to literature, a single training session can significantly decrease the probability of confusion or unfamiliarity of the expected requirements of a test (Rushton, Skuy, & Fridjhon, 2003). The test battery was completed in identical order to control for any nuisance factors, namely first the CMS story reading, then the CDQ, the CMS questions, the TPT, Prospective memory, Suffolk Reading Scale, Prospective memory, SMT, and Cognitive Abilities Test (CAT). Acceptable prompting levels were set out in the in-school manual and adhered to by the administrators. Four test administrators were available to lend assistance to learners who may have needed help during the session. The test battery took about an hour and ten minutes to

complete. Between the TPT and Prospective memory a short break was provided during which learners did not leave their tables but stretching was done. Every child put his/her tests in an envelope on completion and these envelopes were collected by the test administrators at the end of the session. The child packs were marked down on the class lists to ensure that all packs were collected. The learners were then debriefed after the session.

3.6 Threats to Validity

Possible threats to validity during the research, whose exact impact on the results are unclear, included the following:

Firstly, it was impossible to ensure that the test situations were identical. There is a potential for diffusion owing to the dual administrations. In order to minimize the discrepancy between the test situations, adherence to the in-school protocol was vital. Testing sessions took place at the same time. Despite administrator training, it is likely that some administrators were unfamiliar with the tests, but more so with the constructs measured. This may have impacted results, such as in the case of the TPT, where results are sensitive to external distractions, even from the administrators themselves. Next, language problems could have unduly influenced the test results, regardless of the fact that English is the language of instruction. As a result, it is possible that some learners did not understand what was expected of them, even if the sample items were completed. The limited range of English comprehension may have further complicated accurate testing, such as in the case of the CMS, which is a verbal test. Not only may it have been

difficult for some learners to understand the stories, but literature indicates that it is easier to remember familiar items or items that have some meaning (Baddeley, 2004).

The heavier rate of attrition of the CMS was due to latecomers. As this was the first test in the battery, the sample size of completed CMS tests (n=187) was somewhat smaller than for either the TPT (n=205) or the SMT (n=190), complicating data analysis, and possibly influencing external validity of the tests.

Fatigue is another factor that could have unduly influenced test results. The battery took one hour and 10 minutes to complete and for learners of this age, this may have led to weariness. The break given was intended to help counter fatigue and attempts were made to control rather than eliminate this factor by completing the tests in the same order.

Although the administrators did not use the word 'test' when introducing themselves and the study, the time limit on some of the tests was reminiscent of a test situation, and it is possible that some learners experienced anxiety with regard to the testing situation. Also, the situation was new and included unfamiliar test administrators, which could also cause or exacerbate anxiety, possibly unduly influencing the test results. On some of the more difficult items of the SMT, there was evidence of panic, in that some respondents marked nearly every letter. As false positives are subtracted from correct responses during scoring, these results were in fact out of all proportion.

3.7 Data Analysis:

Firstly, a Cronbach coefficient alpha analysis was done to determine the internal consistency reliability estimates for the CMS and a split-half method, by which items were split between even and uneven numbers, was used for the SMT with the purpose of determining the tests' reliability as little psychometric information was available on any of the three tests. An item-by-item analysis was done in order to assess item difficulty, and item variance was used to assess the usefulness of items (Rosnow & Rosenthal, 2008). However, the TPT results only provided totals, not item scores, making internal reliability estimates and an item-by-item analysis impossible.

The suitability for parametric analysis had to be investigated prior to investigation of the research questions. To establish whether the tests were normally distributed, histograms and the Kolmogorov-Smirnov test of normality were done (Salkind, 2004). A series of Pearson Product Moment Correlation Coefficients (r) were conducted to investigate correlations between the three tests, as well as between the male and female groups. A t-test for independent samples was done on the mean scores of the two genders. Fisher-z transformations were used to assess whether differences in the correlations between males and females were significant (Howell, 2008). The findings are presented and discussed in subsequent chapters.

CHAPTER FOUR

RESULTS

4.1 Introduction

A series of statistical techniques was conducted to investigate the research questions. First, the reliability of the tests as a whole and per item was inspected. Next, correlations between the different tests were investigated, also with regards to gender. Lastly, a correlation between correlation matrices of the genders was explored.

4.2 Reliability of the Instruments

It was essential to first confirm the reliability of the three tests (CMS, SMT, and TPT) used and verify their consistencies within the study. Very little psychometric data was found on the original versions of the abovementioned tests. As mentioned before, adaptations of the original tests were used to make them more applicable to the study, for instance changing the CMS to a group test for ease of application. These changes could potentially impact reliability of the tests. Cronbach Coefficient Alphas (α) were calculated for the CMS as measure of internal consistency reliability of the instrument. As the SMT is a timed test in which items progressively become more complicated, a split-half method was used to estimate internal reliability in this test (Kaplan, 1987). Unfortunately no information on the individual items of the TPT was available since the original study focussed on the totals as opposed to the separate items; hence, no internal reliability could be calculated for this test.

Table Two indicates that within this study, the Conceptual Recall sections of the CMS exhibited a particularly strong level of consistency ($\alpha=0.99$; $\alpha=0.97$). However, the Recognition sections had reliability estimates of $\alpha= 0.43$ and $\alpha=0.28$, indicating very low reliability, and suggesting a lack of consistency in its measurement of recognition memory. However, the overall reliability estimate for the CMS was $\alpha=0.72$, signifying that this test measures memory consistently and suggesting that it is reliable for this sample. The SMT had moderately acceptable reliability estimate ($\alpha=0.68$), suggesting that there may be some lack of consistency in its measurement of working memory and attention.

Table 2: Reliability Estimates for Child Memory Scale and the Search and Memory Task

<i>Cronbach Coefficient Alphas</i>		
		<i>Standardized Alpha</i>
CMS:	Conceptual Recall 1	0.999958
	Conceptual Recall 2	0.972588
	Recognition 1	0.429360
	Recognition 2	0.282674
	Information recall (Total)	0.722330
SMT:	LD_DS	0.679629

As already mentioned, item variance was used to assess the usefulness of items (Anastasi, 1988; Rosnow & Rosenthal, 2008). An item-by-item analysis was done to assess item difficulty,

traditionally defined as the proportion of people who answer the item correctly. As illustrated in Table Three, no items seemed to unduly influence alpha on the Conceptual Recall.

Table 3: Item analysis of Child Memory Scale - Conceptual Recall

<i>Cronbach Coefficient Alphas with Delete Variable</i>					
	Story1		Story 2		
<i>Deleted Variable</i>	<i>Correlation with Total</i>	<i>Standardized Alpha</i>	<i>Correlation with Total</i>	<i>Standardized Alpha</i>	<i>Deleted Variable</i>
CR1_Q1A	0.999688	0.999956	0.711668	0.976281	CR2_Q1A
CR1_Q1B	0.999718	0.999955	0.715279	0.976133	CR2_Q1B
CR1_Q1C	0.999700	0.999955	0.894866	0.968601	CR2_Q2A
CR1_Q2A	0.999690	0.999956	0.895135	0.968589	CR2_Q2B
CR1_Q2B	0.999670	0.999956	0.941280	0.966603	CR2_Q3A
CR1_Q3A	0.999843	0.999952	0.941280	0.966603	CR2_Q3B
CR1_Q3B	0.999767	0.999954	0.941279	0.966603	CR2_Q3C
CR1_Q3C	0.999820	0.999952	0.941277	0.966603	CR2_Q4A
CR1_Q4A	0.999900	0.999950	0.941286	0.966602	CR2_Q4B
CR1_Q4B	0.999885	0.999951	-	-	-

As shown in Table Four, items 1, 3, 10, 11, and 14 of the Recognition Memory 1 and Items 2, 4, 5, 6, 8, 15 of the Recognition Memory 2 seemed to be especially problematic with regards to internal reliability on this section of the CMS.

Table 4: Item analysis of Child Memory Scale-Recognition memory

<i>Cronbach Coefficient Alphas with Deleted Variable</i>					
Story 1			Story 2		
<i>Deleted Variable</i>	<i>Correlation with Total</i>	<i>Standardized Alpha</i>	<i>Correlation with Total</i>	<i>Standardized Alpha</i>	<i>Deleted Variable</i>
R1_1	0.027226	0.442513	0.035076	0.286287	R2_1
R1_2	0.143421	0.411720	-0.026267	0.308020	R2_2
R1_3	0.063114	0.433156	0.106466	0.260225	R2_3
R1_4	0.182362	0.401073	-0.012874	0.303326	R2_4
R1_5	0.256320	0.380390	0.001961	0.298094	R2_5
R1_6	0.165182	0.405790	-0.007061	0.301280	R2_6
R1_7	0.332468	0.358451	0.141626	0.247078	R2_7
R1_8	0.109151	0.420953	-0.033399	0.310507	R2_8
R1_9	0.141514	0.412237	0.170598	0.236089	R2_9
R1_10	0.010055	0.446942	0.273837	0.195761	R2_10
R1_11	0.052339	0.435980	0.064472	0.275657	R2_11
R1_12	0.223561	0.389626	0.187762	0.229512	R2_12
R1_13	0.220880	0.390377	0.207518	0.221878	R2_13
R1_14	0.047229	0.437315	0.167532	0.237259	R2_14
R1_15	0.155444	0.408450	-0.000650	0.299018	R2_15

On the SMT (Table Five), question 7 appears to be problematic with regard to internal reliability of the SMT.

Table 5: Item analysis of the Search and Memory Task

Cronbach Coefficient Alphas with Delete Variable

<i>Deleted Variable</i>	<i>Correlation with Total</i>	<i>Standardized Alpha</i>
Ldq1A	0.445002	0.642135
Ldq1B	0.492041	0.635274
Ldq2A	0.438032	0.643144
Ldq2B	0.211442	0.674723
Ldq3A	0.413128	0.646727
Ldq3B	0.297997	0.662929
Ldq4A	0.268482	0.666988
Ldq4B	0.306669	0.661729
Ldq5A	0.211043	0.674776
Ldq5B	0.343409	0.656609
Ldq6A	0.204798	0.675614
Ldq6B	0.257910	0.668432
Ldq7A	0.167775	0.680549
Ldq7B	0.107503	0.688455

4.3 Normality of the data

Certain assumptions are prerequisites for using parametric techniques. These include randomness and independence of sampling and testing; homogeneity of residual variances among groups, and normality of residuals in groups (Howell, 2008; Kaplan, 1987). It was thus vital to ascertain whether test results are normally distributed or not. Measures of central

tendency, histograms and the Kolmogorov – Smirnov test of normality were used to determine normal distribution.

Examination of the histograms for the three tests revealed that the CMS data (Figure 1) were slightly negatively skewed, indicating that the results are approximately normally distributed. Both SMT and TPT data (Figures 2 & 3 respectively) were significantly negatively skewed, signifying that the majority of learners performed extremely well on these tests. These visual examinations were confirmed through examination of the measures of central tendency for the three tests, focusing particularly on the distance between the mean and median (Howell, 2008).

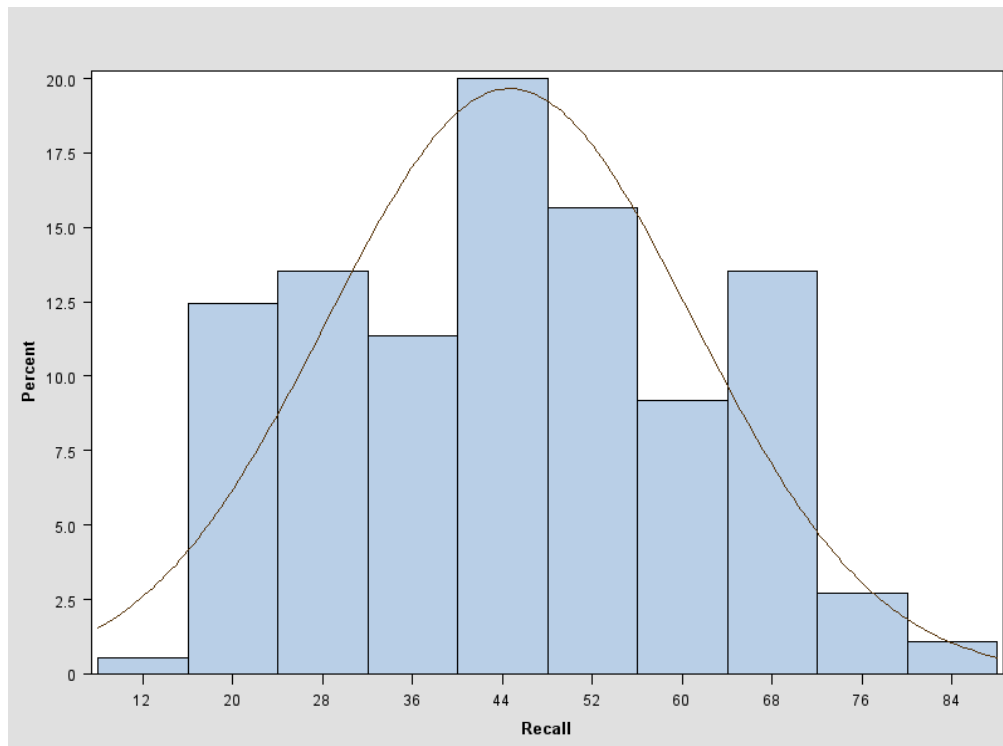


Figure 1: Histogram: Child Memory Scale

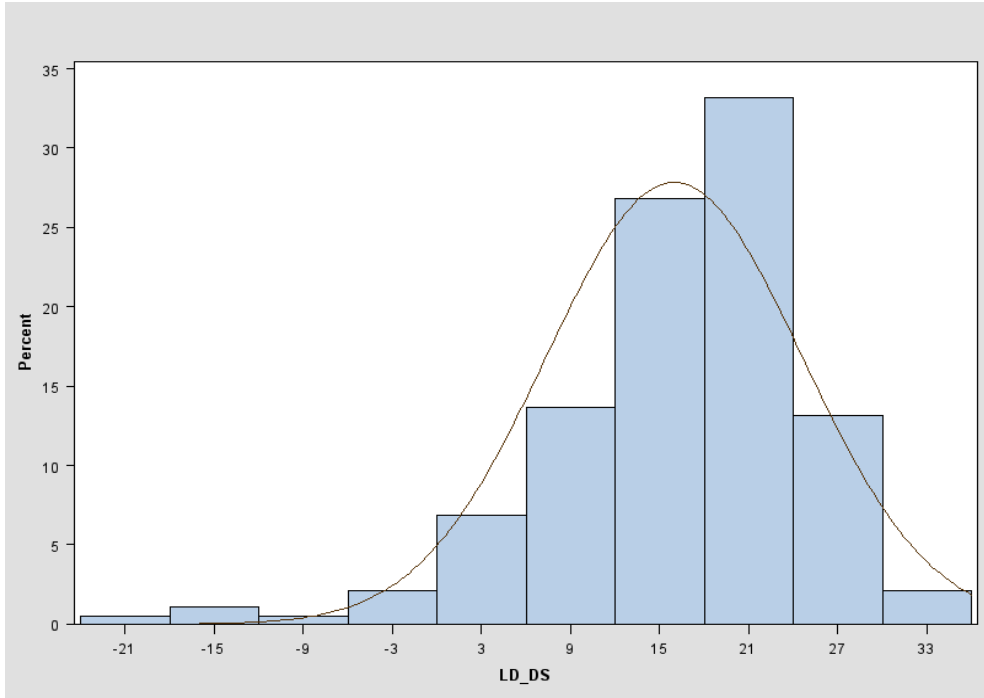


Figure 2: Histogram: Search and Memory Task

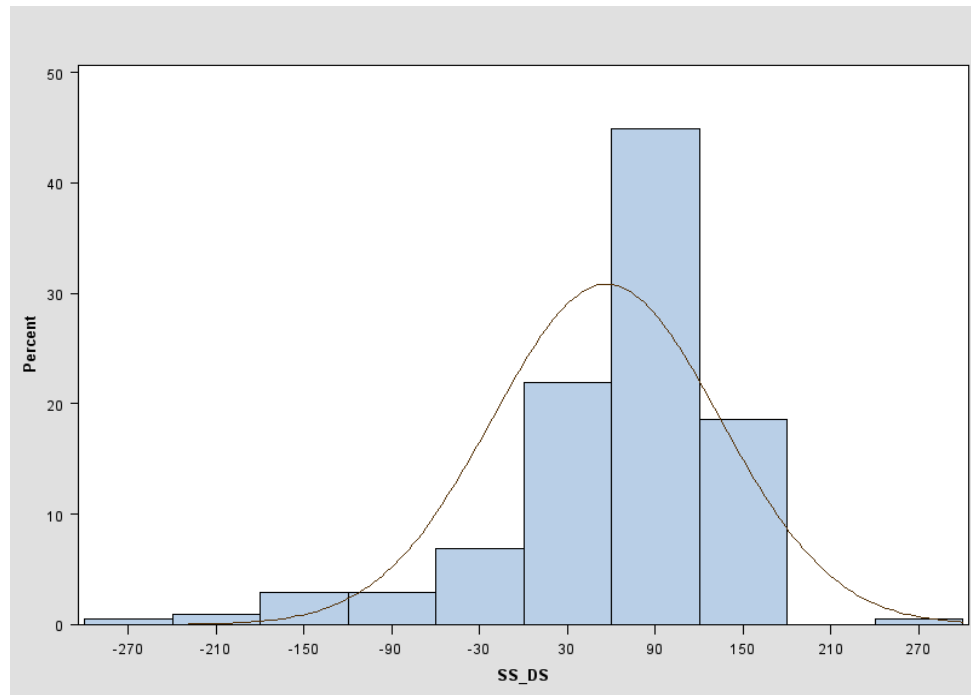


Figure 3: Histogram: Toulouse-Pieron Test

In Table Six it is clear that the distances between the median and the mean of the SMT and the TPT are considerable, confirming that the tests are not normally distributed.

Table 6: Basic Descriptive Statistics for the Tests

Simple Descriptive Statistics

<i>Variable</i>	<i>N</i>	<i>Mean</i>	<i>Std. Dev.</i>	<i>Median</i>	<i>Mode</i>	<i>Minimum</i>	<i>Maximum</i>
CMS	185	44.66486	16.23356	44.0000	43.0000	14	84
SMT	190	15.97368	8.59644	17.0000	17.0000	-20	33
TPT	205	56.37073	77.71289	70.0000	0.0000	-288	281

An evaluation of the normality of the distributions utilizing Kolmogorov-Smirnov Tests of Normality was also done.

According to the results in Table Seven, none of the tests could strictly be considered normally distributed. This calls into question the use of parametric tests to analyze the data. However, the Kolmogorov-Smirnov tests are particularly strict and thus sensitive to inconsistencies (Howell, 2008). According to Kaplan (1987) the probability of normal sample distribution increases as sample size increases. Considering the large sample size of this study, it was decided to persist with parametric analysis, but that it was advisable to conduct parallel non-parametric tests in certain cases as a means of confirmation of parametric results.

Table 7: Kolmogorov-Smirnov Tests of Normality for the Tests

Kolmogorov-Smirnov Goodness to Fit

	<i>Statistic (D)</i>	<i>p-Value</i>
CMS	0.07425	p=0,014
SMT	0.14078	p<0,0100
TPT	0.15765	p<0,0100

*p<.05 **p<.01

4.4 Correlation Matrices Between the Child Memory Scale, Search and Memory Task and Toulouse-Pieron Test

To establish the degrees of correlation between episodic memory and cued recall (CMS), working memory and attention (SMT); and sustained visual attention (TPT), a series of Pearson Product Moment Correlation Coefficients (r) was done to show the strength and direction of the association between the three tests. Pearson Product Moment Correlation Coefficients (r) were also done to show the strength and direction of the association between the two genders on each of the three tests. A t-test for independent samples was done on the mean scores of the two genders on the three tests to examine the measure of the relationship between the two independent variables.

Results, as shown in Table Eight, indicated that a relatively weak, positive but insignificant correlation existed between the SMT and the CMS ($r = 0.0557$ and $p > 0.05$; $r = 0.07236$ and $p > 0.05$). They also pointed to weak, positive relationships existing between males' scores on the CMS and the SMT ($r = 0.05092$; $p > 0.05$) and between females' scores on the same two tests ($r = 0.07384$; $p > 0.05$).

Table 8: Relationship between the Child Memory Scale and the Search and Memory Task by Gender

<i>Pearson and Spearman Correlation Coefficients</i>					
Overall: N=181	CMS		Male n=84	CMS	
	Pearson's	Spearman's		Pearson's	Spearman's
	CMS			CMS	0.05092
				SMT	
		0.0557			0.07236
	SMT		Female n=95	CMS	0.07384
				SMT	

* $p < .05$ ** $p < .01$

Results in Table Nine indicated a significant, weak, and positive correlation ($r = 0.20824$ and $p < 0.01$; $r = 0.29366$ and $p < 0.0001$) between the TPT and the CMS. There are also indications of weak, positive correlations between males' scores on the CMS and the TPT ($r = 0.1552$; $p > 0.05$), as well as between females' scores ($r = 0.22564$; $p > 0.05$) on the same two tests.

Table 9: Relationship between the Child Memory Scale and the Toulouse-Pieron Test by Gender

<i>Pearson and Spearman Correlation Coefficients</i>					
Overall: N=184	CMS		Male n=84	CMS	
	Pearson's	Spearman's		Pearson's	TPT
	CMS			CMS	0.15852
		0.20824**		TPT	
	TPT		Female n=95	CMS	0.22564
		0.29366**		TPT	

*p<.05 **p<.01

Both the Pearson's and the Spearman's correlations (Table Ten) indicated a weak positive correlation between the TPT and the SMT ($r = 0.11572$; $\rho = 0.19667$), but the Pearson's indicated an insignificant correlation ($p > 0.05$), whereas the Spearman's points to a significant correlation ($p < 0.01$). An examination of individual scores on the TPT and the SMT reveals several outliers, and as Spearman's Correlations tend to be less effected by such extreme scores (Howell, 2008), accepting the correlation as significant seems to be in order.

Males' scores on the TPT and the SMT show a weak, positive relationship ($r = 0.10057$; $p > 0.05$) as did females' scores on the same tests ($\rho = 0.15743$; $p > 0.05$). Since none of the correlations proved to be significant, the null hypothesis () could not be rejected, indicating

that there were no significant differences between males' and females' performances on the three tests.

Table 10: Relationship between the Search and Memory Task and the Toulouse-Pieron Test by Gender

Pearson and Spearman Correlation Coefficients

	TPT		TPT	
	Pearson's	<i>Spearman's</i>		<i>Pearson's</i>
Overall: N=189	TPT		Male n=84	TPT 0.10057
		0.11572		SMT
		0.19667**	Female n=101	TPT 0.15743
	SMT			SMT

*p<.05 **p<.01

As presented in Table 11, the results point to probabilities of 0.0184, 0.1909 and 0.2033 for CMS, SMT and TPT, respectively. This indicates that the null hypothesis () could not be rejected, suggesting no significant difference between the scores of the two genders.

Table 11: Relationships between the Child Memory Scale, the Search and Memory Task and the Toulouse-Pieron Test by Gender

T-test for Independent Samples

		<i>N</i>	<i>F</i>	<i>t</i>	<i>Mean</i>	<i>Degrees of Freedom</i>
CMS	<i>Male</i>	85	1.09	2.38	48.0833	178
	<i>Female</i>	96			42.4063	
SMT	<i>Male</i>	85	1.50	-1.31	15.2588	184
	<i>Female</i>	101			16.9010	
TPT	<i>Male</i>	85	1.11	1.28	68.4588	185
	<i>female</i>	102			53.6275	

*p<.05 **p<.01

4.5 Correlation between the gender correlation matrices

Fisher–z transformations were done on the gender matrices to examine if there is a significant correlation between them.

Results (Table 12) indicate z-scores of -0.15053 for SMT/CMS ($p>0.05$), -0.45727 for CMS/TPT ($p>0.05$), and -0,3852 for TPT/SMT($p>0.05$). None of these were significant, suggesting no significant difference in the correlation between the gender matrices and indicating measurements were similar between genders.

Table 12: Fisher-Z transformations on gender correlations between CMS, SMT and TPT

Fisher-Z Transformations (z)

CMS/SMT	-0.15053
CMS/TPT	-0.45727
TPT/SMT	-0.3852

*p<.05 **p<.01

CHAPTER FIVE

DISCUSSION

5.1 Introduction

The focus of this research was to contribute to the body of psychometric knowledge of testing and assessment in the South African context, particularly with regards to the Child Memory Scale, the Search and Memory Task, and the Toulouse-Pieron Test as measures of memory. Although these tests have been used in international studies (Matsui et al., 2004; Stansfeld et al., 2005), their psychometric properties in a South African context have not yet been researched. Given the importance placed by legislation and the HPCSA on fair and unbiased testing, empirical research concerning the applicability of psychological measuring instruments for South African populations, appears imperative. The tests' internal reliabilities, the inter-correlations between the tests as well as between the performances of the two genders on the tests, and as such their construct validities, were explored and evaluated.

5.2 Discussion of Results

5.2.1 Internal reliability

In terms of exploring internal reliability of the CMS and the SMT, Cronbach Coefficient Alphas and the split-half method were employed, respectively. The CMS is a test for auditory episodic

memory and can be divided into Conceptual Recall and Recognition sections. The Conceptual Recall sections had particularly high alphas and none of the items appear to negatively influence the overall internal reliability of these sections. Corresponding to the particularly low alphas of the Recognition sections, Items 1, 3, 10, 11, and 14 on Recognition Memory 1 and Items 2, 4, 5, 6, 8, and 15 on Recognition Memory 2 proved to be unduly influencing these sections' internal reliability on the whole. In fact, all the Recognition Memory sections were problematic in terms of the CMS's overall internal reliability. The discrepancy between these two parts of the test was an unexpected result. The Conceptual Recall sections required the learners to answer questions in written form while the Recognition parts of the test required the learners to simply tick 'yes' or 'no'. Considering the fact that English is a second language to many of the learners, the Recognition parts appeared to be 'easier' to answer. However, the Conceptual Recall section consisted of questions that were in written form on the answer sheets for the learners to read. For the duration of the Recognition section, learners had to listen to questions asked by an administrator without the benefit of seeing the questions on paper. Given the language difficulties that some of the learners experience, it is possible that the fact that the learners were unable to read the questions to facilitate understanding may have unduly influenced the results. In addition, whilst answering the Conceptual Recall sections, the learners could answer the questions at their leisure. As the questions for the Recognition sections were read by the administrator, it is possible that some learners felt pushed to answer as soon as possible, leading to a lack of consistency in the measurement of recognition memory. By providing learners with the questions for the Recognition sections, the internal reliability of this part of the CMS may be improved. Still, the overall internal reliability of the CMS appeared to be acceptable in this study.

The SMT's internal reliability was lower but acceptable, suggesting a degree of inconsistency in its measurement of working memory and attention. The SMT is a timed test in which items become progressively more difficult. These two traits of the test complicated the computation of an estimation of the internal reliability of the instrument, necessitating the use of the split-half method. The whole of Item 7 proved to be problematic with regards to internal reliability. Given the fact that items on the SMT increasingly become more complicated leading to higher anxiety, this finding is not surprising. However, the inconsistency of the reliability of Item 7 is perturbing, and as such the item could be discarded to increase the internal reliability of the SMT.

It was not possible to estimate the internal reliability of the TPT due to lack of information regarding the individual items of the test.

Neither the results of the SMT nor the TPT were normally distributed. It is striking that both these tests are nonverbal, possibly explaining the tendency of learners to do very well on these tests.

Several factors may have influenced the learners' performances on the three tests. The completion of the CMS involves a fair amount of language abilities. Although the language of instruction is English, during testing it was clear that some of the learners had difficulty understanding English instructions. Language is bound to have influenced the results unduly. Since more than half of the administrators were Caucasian, learners may have refrained from asking for help if they did not understand the questions, as most learners were either of Zulu or

Indian origin. As already discussed, it is possible that due to limited language abilities, learners did better on questions that were visually represented, as opposed to the auditory items.

The TPT measures sustained attention, whereas the SMT measures working memory and attention. Success on these tests was in part dependent on good visual scanning and psychomotor abilities. Two groups of administrators collected the data, and there is bound to have been differences in their behaviour during the tests in the different groups. Attention tests are very sensitive to distractions of any kind, including test administrators moving around in the classroom. Test fatigue (Anastasi, 1988), to which attention tests are sensitive to as well, may have played a role in the scores as the TPT was administered 3rd and the SMT 5th in a list of six tests. Considering the socio-economic circumstances of a lot of the learners, low blood sugar level as a result of hunger could have negatively impacted the test scores. HIV status may also have affected learners' performance on the test, since positive HIV status has been linked to motor difficulties, nutritional problems due to health complications, slow reaching of milestones, poor school performance (NIAID, 2009), as well as vision loss, nerve damage, brain impairment and loss of co-ordination in advance stages of the illness (Epigee, n.d.). Considering that KwaZulu-Natal is currently the province with the highest prevalence of HIV in the general population, estimated at 15.8% by the South African National HIV Survey in 2008 to 18.8% of adults infected (Actuarial Society in South Africa (ASSA), 2008), reflection on HIV as a potential source of bias seems imperative, even if the actual number of HIV positive learners in the current study is unknown. Since the answer sheet consisted of small figures between which the learners had to discriminate, a learner with visual perception or discrimination difficulties would have been unfairly disadvantaged. Furthermore, malnutrition affects both the physical and mental

state and may also have an effect on the test results. Symptoms may be difficult to identify by others, such as teachers, unless the malnutrition is severe, and may include chronic fatigue, slowed reflexes, inattentiveness and memory troubles, muscle weakness or trembling and visual difficulties because of a lack of vitamin A (Alward, 2005). There is also a possibility that different language groups have different response styles to timed tests, such as the TPT and SMT. Differing response styles have been found to exist between cultures (Chen, Lee, & Stevenson, 1995) on rating scales. However, little information exists on the response styles of Zulu and Indian learners in South Africa and it is unclear how they would differ, except to hypothesize that Zulu learners may have worked slower and taken more time to make a decision, whereas Indian learners may have worked faster, but made more mistakes. High anxiety during timed tests is common and can adversely influence test scores (Anastasi, 1988), and was evident in some learners' response sheets in that they marked as many of the figures and target letters as they could, working to their disadvantage as the scoring of the TPT and the SMT penalize a respondent for 'false positives'. As the school is situated 604 meters from the flight path of the local airport, the noise levels of the ascending and descending airplane may have been a factor in the testing. Eighteen flights were registered during the administration of the test battery with a maximum noise level of 91.8 (LaMax) and an average of 63.5 (Leq), placing the school in a moderate to high noise range. This may have interfered with the measurements of some of the tests, but there is no way to ascertain during which tests specific flights were taking off or landing, making it difficult to ascribe noise as a source of bias in any particular test.

5.2.2 Construct Validity

In terms of the construct validity of the tests (Cronbach & Meehl, 1955; Salkind, 2004), two sets of correlations as well as Fisher-z transformations, were conducted. The first set was to establish the nature of correlations between the three tests and the second, on gender differences between the different tests. The Fisher-z transformations were aimed at ascertaining if significant correlations between the different gender matrices existed.

The research results indicated a relatively weak but significant relationship between the nonverbal TPT that measures sustained attention and the verbal CMS that measures long-term episodic memory and cued recall. This suggests that there is a measure of overlap between the constructs that the two tests measured within the sample. Better sustained attention is correlated with enhanced long-term episodic and semantic memory. However, according to literature (Lieberman, 2004), episodic memory is in part dependent on attention and the expectation would be a stronger relationship between these constructs. On the CMS, language and administrator bias may have played a major role. Test fatigue (as the CMS was the first test to be administered and the TPT was 3rd), distractions, random marking of items and physical limitations (such as discussed above), may have biased results on the TPT.

A relationship between the TPT and the SMT was to be expected as sustained attention (TPT), and working memory and attention (SMT) are related. However, results indicate a significant but weak relationship. The SMT measures working memory and attention and as such may have been vulnerable to similar influences to the TPT. The outlay of the test may have unduly influenced the results. According to literature, individuals tend to be more successful in

remembering initial and most recent detail, as opposed to ‘middle’ detail. These are called the primacy and the recency effects, respectively (Bower, 2000). As the SMT target letters were presented to the learners in a line, it is conceivable that the learners may have tended to remember initial and ending letters better than others, influencing the amount of learners that would correctly identify these letters as specific items. When considering the SMT answer sheet and how the items are visually displayed on the page, it is also possible that certain items were ‘easier’ to identify due to their placement on the page as in each row there was one target letter apiece in the first, second and third part of that line. Timed tests invoke anxiety in respondents due to the time limit, leading to anxiety. For some, the pressure may have led them to panic and mark as many letters as possible. This worked to their disadvantage as they were penalized for ‘false positives’, as in the TPT. Also, the SMT was the second last test to be administered, making test fatigue as a nuisance factor very probable and maybe even more of a factor than in the case of TPT. Since attention and working memory tests are sensitive to physical conditions such as fatigue, low blood sugar and any other disease that may have an adverse affect on a learner’s ability to concentrate and their psycho-motor speed, results may have been unduly influenced. Distraction from different sources, including the test administrators, may also have impacted on the scores. While both the TPT and the SMT were chosen for the study partly because they were considered ‘non-verbal’ and in order to limit language bias, it is doubtful that language had no effect on the scores. Simply conducting the tests required a degree of understanding of the English language. Communication problems and/or ambiguous directions may have confused some learners, especially in the context of language difficulties.

No significant relationship was found between the CMS and the SMT. These tests measure long-term and short-term memory, respectively and the lack of correlation between them seems out of the ordinary. Atkinson and Shrifin (1968) suggested that information is transferred from the sensory memory to short-term memory (if attention is given) to the long-term memory, mainly through repetition, while other researchers such as Gathercole and Martin (1996) and Logie (1999) suggested that long-memory is involved in short-term memory almost from the very beginning of the process. There thus appears to be a connection of sorts between these two types of memory, which is not reflected in these results. Test fatigue, for which the SMT is very sensitive, may account in part for the lack of correlation, considering that the CMS was the first and the SMT the second last test to be completed in the test battery. Also, high anxiety was, as previously mentioned, may have further influenced the results of the SMT, as it is a timed test.

The results indicate that there were no significant differences in the scores of male and female learners between any of the tests. There was also no significant difference in the correlation matrices of the genders. There is little agreement in literature on the occurrence and nature of differences in cognitive abilities between the genders, as discussed previously. However, it is possible that several factors may have mediated each other, 'disguising' any differences between the genders (Halpern, 1986). More research is needed to explore potential factors and their interaction.

5.3 Limitations of the study

Since this study was an archival study, the researcher had no control over testing circumstances.

The researcher was unable to examine actual answer sheets for patterns, such as different response styles with regard to gender or race. Furthermore, no individual item scores on the TPT were available, making it impossible to calculate reliability estimates for this test.

It was impossible to separate the effects of home language, race, and socio-economic status. This has serious implications in the applicability of the study since it is impossible to assign the reasons for the results. The interaction between factors was not possible to assess, and some factors may have mediated others.

Internal validity may have been threatened by several factors, such as personal history, subject related artifacts, experimenter-related artifacts, demand characteristics, and the experimenter expectancy effect (Rosnow & Rosenthal, 2008).

The adapted SMT was problematic, producing skewed results and making estimation of internal reliability difficult.

Lack of an effective communication language complicated the data collection immeasurably, as some learners could barely understand the instructions to some of the tests. This is a common problem in South Africa, once again indicating test results need to be interpreted with the utmost care.

The effect of the test administrators may have been considerable. Even if the data collection happened at the same time in the different groups, and despite the protocol booklet, some administrators may have digressed. Their individual personality traits, different levels of expertise, and race may have had an impact on whether the learners felt at ease during the testing and if they felt confident enough to ask for help if they needed it.

5.4 Contribution to Knowledge

Regardless of the abovementioned limitations, some positive contributions stemmed from this research. No research has ever been done on these tests' psychometric properties in the South African context, and both the CMS and the SMT, with minor changes, appear to have acceptable internal reliability for this sample of learners. However, the lack of expected strong correlations between the three memory tests, is disturbing. This suggests that the construct validities of these tests are suspicious for this sample of learners. It would appear that several possible factors may have impacted the results on these tests and acted as sources of test bias (Foxcroft & Roodt, 2005). This underscores the importance of not making assumptions about the reliability and validity of measuring instruments for different cultural groups but to confirm their applicability before usage. Furthermore, it highlights the necessity for further research regarding potential sources of bias in South African populations.

5.5 Directions for Future Research

This study, while not definitive, provides numerous possibilities of further aspects to be studied. Future research may focus on the possibility of different response styles of different cultural groups to specific tests. It is possible that particular response styles are more beneficial than others on certain tests. For instance, a child who marks many different target items quickly, may make more mistakes, but may end up with a better score, whereas a child that works slower and considers his/her choices, may make less mistakes but may have difficulty finishing the questions, negatively impacting his/her scores. The implications of testing in English as language of instruction may be well worth exploring, and specifically the degree of bias may be of great interests for other researchers. The identification of other sources of bias for psychometric assessment that are specific to the South African context, such as the extent to which HIV influences psychological test results, may be invaluable. Replications of the current study are imperative to confirm or refute the finding in this report. The RANCH-SA study involved five different schools and it may be worth the time and effort to ascertain if similar patterns of distribution of results exist in the other schools that were not part of this study, and the implications of such findings. Even after all this time, in literature there is little consensus on, for instance gender differences on cognitive abilities (Halpern, 1986). There is also little research concerning the interaction between factors such as language difficulties, race, age and gender, and how they mediate each other.

5.6 Conclusion

In South Africa, research on fair and unbiased psychometric testing for all different populations is still in its infancy and the sheer amount of work that is still to be done, can be daunting. When used appropriately, psychometric testing can contribute to various processes, such as organizations, schools and even the affirmative action process. This study represents only a preliminary and exploratory investigation, another step in the process towards sensitive and impartial assessment for all cultural groups.

References

- Actuarial Society in South Africa. (2008). *Health Statistics: HIV prevalence (%) (total) population*. Retrieved November 15, 2009, from <http://www.healthlink.org.za/healthstats/84/data>
- Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology, 87*, 85–106.
- Alward, M. M. (2005, September 30). *Children and hunger and Malnutrition*. Retrieved November 15, 2009, from <http://wwwgoogobits.com/articles/p3-2824-children-and-hunger-and-malnutrition.html>
- American Educational Research Association. (1999). *Standards for educational and psychological tests*. Washington, D. C.: Author.
- Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan.
- Atkinson, R.C. & Shiffrin, R.M. (1968). Human memory: a proposed system and its control processes. In K.W. Spence (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory, Vol. 2* (pp. 89–195). New York: Academic Press.
- Awh, E., Jonides, J., Smith, E. E., Schumacher, E. H., Koeppel, R. A., Katz, S. (1996). Dissociation of storage and rehearsal in verbal working memory: Evidence from positron emission and tomography. *Psychological Science, 7*(1), 25-31.
- Baddeley, A. D. (1986). *Working memory*. London: Oxford University Press.
- Baddeley, A. D. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology, 49A*(1), 5-28.

- Baddeley, A. (1997). *Human memory Theory and practice*. Revised Edition. UK: Psychology Press.
- Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417-423.
- Baddeley, A. D. (2002). Is working memory still working? *European Psychologist*, 7(2), 85-97.
- Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews/Neuroscience*, 4, 829-839.
- Baddeley, A. D. (2004). The psychology of memory. In A. D. Baddeley, M. D. Kopelman and B. A. Wilson (Eds.), *The Essential Handbook of Memory Disorders for Clinicians* (pp. 1-13). West Sussex, UK: John Wiley & Sons, Ltd.
- Baddeley, A. D., Gathercole. S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105, 158-173.
- Baddeley, A. D. & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *The Psychology of Learning and Motivation*, Vol. 8 (pp. 47-90). New York: Academic Press.
- Biesheuvel, S. (1949). Psychological tests and their application to non-European peoples. In G. B. Jeffrey (Ed.), *The Yearbook of Education* (pp.49-63). London: Evans.
- Biesheuvel, S. (1959). The development of personality in African cultures. Confidential report, Johannesburg: National Institute for Personnel Research
- Bower, G. H. (2000). A brief history of memory research. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford Handbook of Memory* (pp. 3-32). New York: Oxford University Press.
- Broadbent, D. E. (1958). *Perception and communication*. New York: Oxford University Press.

- Campbell D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Chen, C., Lee, S., & Stevenson, H.W. (1995). Response styles and cross-cultural comparisons of rating scales among East-Asia and North American students. *Psychological Science*, 6(3), 170-175.
- Clark, C. R., Moores, K.A., Lewis, A., Weber, D. L., Fitzgibbon, S., Greenblatt, R., et al. (2001). Cortical network dynamics during verbal working memory function. *International Journal of Psychophysiology*, 42, 61-176.
- Cohen, M. J. (1997). *Children's Memory Scale Manual*. San Antonio, Texas: The Psychological Corporation Harcourt Brace and Company.
- Cole, N. S. & Moss, P. A. (1989). Bias in Test Use. In R. L. Linn (Ed.). *Educational Measurement* (3rd ed., pp. 201-219). New York: Macmillian Publishing Company.
- Cronbach, L.J. & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302.
- Darling, S., Della Sala, S., Logie, R. H., & Cantagallo, A. (2006). Neuropsychological evidence for separating components of visuo-spatial working memory. *Journal of Neurology*, 253(2), 176-180.
- Epigee (n.d.). *HIV Symptoms*. Retrieved November 15, 2009, from http://mmm.epigee.org/health/hiv_symptoms.html
- Eysenck, M. W. & Keane, M. T. (2000). *Cognitive psychology* (4th ed). Philadelphia, PA: Psychology Press.

- Fick, M. L. (1929). Intelligence test results of poor white, native (Zulu), coloured and Indian school children and the educational and social implications. *South African Journal of Science*, 26, 67-81.
- Fick, M. L. (1939). *The educability of the South African native*. South African Council for Educational and Social research; Research Series No. 8: Pretoria: South African Council for Educational and Social Research.
- Foxcroft, C. (2002). Ethical issues related to psychological testing in Africa: What I have learned (so far). In W. J. Lonner, D. L. Dinnel, S. A. Hayes, & D. N. Sattler (Eds.), *Online Readings in Psychology and Culture*. Retrieved July 21, 2009, from <http://www.ac.wvu.edu/~culture/foxcroft.htm>
- Foxcroft, C. & Roodt, G. (Eds.) (2005). *An Introduction to Psychological Assessment in the South African Context* (2nd ed.). Cape Town: Oxford University Press.
- Foxcroft, C. D., Roodt, G., & Abrahams, F. (2005). The practice of psychological assessment: Controlling the use of measures, competing values, and ethical practice standards. In C. D. Foxcroft & G. Roodt (Eds.), *An introduction to psychological assessment in the South African context* (2nd ed., pp. 69-82). Cape Town: Oxford University Press.
- Garavan, H., Ross, T. J., Li, S. J., & Stein, E. A. (2000). A parametric manipulation of central executive functioning. *Cerebral Cortex*, 10, 585-92.
- Gathercole, S. E. & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: a longitudinal study. *Journal of Memory and Language*, 28, 200-213.
- Gathercole, S. E., & Baddeley, A.D. (1990a). The role of phonological memory in vocabulary acquisition: a study of young children learning new names. *British Journal of Psychology*, 81, 439-454.

- Gathercole, S. E., & Baddeley, A. D. (1990b). Phonological memory deficits in language-disordered children: Is there a causal connection? *Journal of Memory and Language*, 29, 336 – 360.
- Gathercole, S. E. & Hitch, G. J. (1993). Developmental changes in short-term memory: a revised working memory perspective. In A.F. Collins, S.E. Gathercole, M.A. Conway, & P. E. Morris (Eds.), *Theories of memory* (pp.189-209). Hove: Lawrence Erlbaum Associates Publishers
- Gathercole, S. E., & Martin, A. J. (1996). Interactive processes in phonological memory. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp, 73-100). Hove: Psychology Press
- Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The Structure of Working Memory From 4 to 15 Years of Age. *Developmental Psychology*, 40(2), 177–190.
- Giambra, L.M. & Quilter, R. E. (1989). Sex differences in sustained attention across the adult life span. *Journal of Applied Psychology*, 74(1), 91-95.
- Gomes, L.M., Martinho, A.J., & Castelo, N.A. (1999). Effects of occupational exposure to low frequency noise on cognition. *Aviation, Space and Environmental Medicine*, 70(2), A115-A118.
- Gould S. J. (1987). *The mismeasure of man*. UK: Penguin Books.
- Greenfield, P. M. (1997). You can't take it with you: Why ability assessments don't cross cultures. *American Psychologist*, 52(10), 1115-1114.
- Gregory, R. J. (2007). *Psychological testing: History, Principles and Applications* (5th ed.). Wheaton: Pearson Education Inc.
- Halpern, D. F. (1986). *Sex differences in cognitive abilities*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

- Health Professional Council of South Africa (2006, September 20). *F207 List of classified tests*. Retrieved March 20, 2007, from <http://www.hpcs.co.za/hpcs/UserFiles/File/PSYCHOLOGY/F207%20update%20september%202006.pdf>
- Hill, R. D., Grut, M., Wahlin, A., Herlitz, A., Winblad, B., & Backman, L. (1995). Predicting memory performance in optimally healthy very old adults. *Journal of Mental Health and Aging, 1*, 55-65.
- Hoffman, J. E. (1998). Visual attention and eye movements. In H. Pashler (Ed.), *Attention* (pp. 119-153). Hove, UK: Psychology Press Ltd.
- Hofstede, G. (n.d). *Culture*. Retrieved November 25, 2010, from <http://www.geerthofstede.nl/culture.aspx>
- Howell, D. C. (2008). *Fundamental statistics for the behavioural sciences (6th ed.)*. Canada: Thomson Wadworth.
- Hultsch, D. F., Masson, M. E. J., & Small, B. J. (1991). Adult age differences in direct and indirect tests of memory. *Journal of Gerontology: Psychological Sciences, 46*, 22-30.
- Huysamen, G. K. (1980). *Psychological testing in South Africa*. Johannesburg: H & R – Academica.
- Huysamen, G. K. (1994). *Metodologie vir die sosiale en gedagswetenskappe*. Pretoria, RSA: Southern Boekuitgewers.
- Hygge, S., Boman, E., & Enmarker, I. (2003). The effects of road traffic noise and meaningful irrelevant speech on different memory systems. *Scandinavian Journal of Psychology, 44*, 13-21.
- Jones, G., Gobet, F., & Pine, J. M. (2005). Modelling vocabulary acquisition: An explanation of the link between the phonological loop and long-term memory. *Journal of Artificial Intelligence and Simulation of Behaviour, 1*, 509-522.

- Jones, G., Gobet, F., & Pine, J.M. (2007). Linking working memory and long-term memory: A computational model of the learning of new words. *Developmental Science*, 10(6), 853-873.
- Jonides, J., Smith, E. E., Koeppe, R. A., Awh, E., Monoshima, S., & Mintun, M. A. (1993). Spatial working memory in humans as revealed by PET. *Nature*, 363, 623-625.
- Kahneman, D. (1973). *Attention and effort*. New York: Prentice-Hall.
- Kaplan, R. M. (1987). *Basic statistics for the behavioural sciences*. Newton, MA: Allyn and Bacon.
- Klineberg, O. (1980). Historical perspectives: cross-cultural psychology before 1960. In H. C. Triandis & W. W. Lambert (Eds.), *Handbook of Cross-cultural psychology*, Vol. 1 Perspectives. Boston: Allyn & Bacon Inc.
- LaBerge, D. (1983). Spatial extent of attention to letters and words, *Journal of Experimental Psychology: Human Perception and Performance*, 9(3), 371-379.
- Lewin, C., Wolgers, G., & Herlitz, A. (2001). Sex differences favouring women in verbal and non-verbal , but not in visuospatial episodic memory, *Neuropsychology*, 15, 165-173.
- Lezak. M. (1983). *Neuropsychological assessment*. Oxford: Oxford University Press
- Lieberman, D. A. (2004). *Learning and Memory: An integrative Approach*. Toronto, Canada: Wadsworth.
- Logie, R. H. (1995). *Visuo-spatial working memory*. East Sussex: Lawrence Erlbaum Associates Publishers.
- Logie, R. H. (1999). State of the art: Working memory. *The Psychologist*, 12, 174-178.

- López, B., & Vázquez, C. (2003). Effects of the Attention Process Training (APT) on attentional improvement perception in people diagnosed of schizophrenia. *Psychiatry Research*, 119, 41-53.
- Loring-Meier, S. & Halpern , D. F. (1999). Sex differences in visuospatial working memory: Components of cognitive processing. *Psychonomic Bulletin & Review*, 6(3), 464-471.
- Martin, A.L. (1915). Experiments with Binet-Simon tests upon African children. *Training School bulletin*, 12, 122-123.
- Marshall, J. C. & Halligan, P. W. (1994). The yin and the yang of visual-spatial neglect: a case study. *Neuropsychologia*, 32(9), 1037-1057.
- Matsui, T., Stansfeld, S., Haines, M., & Head, J. (2004). Children's cognition and aircraft noise exposure at home - The West London schools study. *Noise & Health*, 7, 25, 49-58.
- Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behaviour*. New York: Holt, Reinhart & Winston.
- Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology*, 7, 480-494.
- Nell, V. (1999). Standardising the WAIS-III and the WMS-III for South Africa: Legislative, Psychometric, and Policy Issues. *South African Journal of Psychology*, 29(3), 128-137.
- National Institute of Allergy and Infectious Diseases (2004, July). HIV Infection in Infants and Children. Retrieved on November 15, 2009, from <http://www.niaid.nih.gov/factsheets/hivchildren.htm>
- O'Shaughnessy, E. (2004). W. R. Bion's theory of thinking and new techniques in child analysis. In E. B. Spillius (Ed.), *Melanie Klein Today: developments in theory and practice Vol. 2*, (pp. 177-189). NY: Brunner-Routledge.

Pashler, H. (1998). Introduction. In H. Pashler, (Ed.), *Attention* (pp.1-11). Hove, UK: Psychology Press Ltd.

Posner, M. I. (1980). Orientating of attention. *Quarterly Journal of Experimental Psychology*, 32, 3-25.

Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, 109(2), 160-174.

Retief, A. (1988). *Method and Theory in Cross-cultural Psychological Assessment*. Pretoria: Human Sciences Research Council.

Rich, S. G. (1917). Binet-Simon tests on Zulus. *South African Journal of Science*, 14, 47-482.

Roshan Cultural Heritage Institute (2001). *Definition of culture*. Retrieved November 1, 2010, from <http://www.roshan-institute.org/templates/System/details.asp?id=39783&PID=474552>

Rosnow, R. L. & Rosenthal, R. (2008). *Beginning behavioural research. A conceptual primer (6th ed.)*. New Jersey: Pearson Education Inc.

Rushton, J.P., Skuy, M. & Fridjhon, P. (2003). Performance on Raven's Advanced Progressive Matrices by African, East Indian, and White Engineering Students in South Africa. *Intelligence*, 31, 123-137.

Salkind, N. J. (2004). *Statistics for people who (think they) hate statistics*. Thousand Oaks, California: Sage Publications Inc.

Schacter, D. L. & Tulving, E. (1994). *Memory Systems 1994*. USA :Bradford Books.

Schein, E. H. (1997). *Organizational Culture and Leadership* (1997, October). Retrieved November 1, 2010 from <http://www.tnellen.com/ed/tc/schein.html>

- Sehlapelo, M. & Terre Blanche, M. (1996). Psychometric Testing in South Africa: Views from Above and Below. *Psychology in Society*, 21, 49-59.
- Shallice, T. & Warrington, E. K. (1970). Independent functioning of verbal memory stores: a neuropsychological study. *Quarterly Journal of Experimental Psychology*, 22(2), 261-273.
- Smith, E. E., Jonides, J., & Koeppel, R. A. (1996). Dissociating Verbal and Spatial Working Memory using PET. *Cerebral Cortex*, 6(1), 11-20.
- Smith, A. P., & Miles, C. (1987). The combined effects of occupational health hazards: an experimental investigation of the effects of noise, nightwork and meals. *International Archives of Occupational and Environmental Health*, 59, 83-89.
- Stansfeld, S.A., Berglund, B., Clark, C., Lopez-Barrio, I., Fischer, P., Öhrström, E., et al. (2005). Aircraft and road traffic noise on children's cognition and health: a cross-national study. *Lancet*, 365, 1942-1949.
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organisation of memory*. London: Academic Press.
- Vallar, G., & Papagno, C. (2002). Neuropsychological impairments of verbal short-term memory. In M. Kopelman, A. D. Baddeley, & B. Wilson (Eds.), *Handbook of neuropsychology* (2nd ed.) (pp. 249-270). Chichester, England: Wiley
- Van de Vijver, F. J. R. & Poortinga, Y. H. (1997). Towards an integrated analysis of bias in cross-cultural assessment. *European Journal of Psychological Assessment*, 13(1), 29-37.
- Van de Vijver F. & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *Revue européenne de psychologie appliquée*, 54, 119-135.

Van den Berg, C. G. (1938). *A comparative study of European, Indian and Zulu school children in Natal as regards intelligence and learning and memory (An investigation based on the results of the Army Beta test and three memory and learning tests)*. Unpublished M.Ed. Thesis, University of South Africa, Pretoria.

Veroff, J. B. & Goldberger, N. R. (1995). What's in a Name? The Case for 'Intercultural' Psychology. In J. B. Veroff & N. R. Goldberger (Eds.), *The Culture and Psychology Reader* (pp.1-15). New York: New York City Press.

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1997). Toward a theory of episodic memory: The frontal lobes and autonoetic consciousness. *Psychological Bulletin*, 121, 331-354.

Yantis, S. (1998). Control of visual attention. In H. Pashler (Ed.), *Attention* (pp.223-256). Hove, UK: Psychology Press.

Appendix A: Parental/guardian consent letter



Psychology

School of Human & Community Development
University of the Witwatersrand
Private Bag 3, WITS, 2050
Tel: (011) 717 4500 Fax: (011) 717 4559



Dear Legal Guardian

My name is Riza van der Merwe and I am conducting research for the purposes of obtaining a Degree in Masters of Educational Psychology at the University of the Witwatersrand. The title of my research is “Application of the Child Memory Scale, the Search and Memory Task, and the Toulouse-Pieron Test to primary school learners”. This study intends to determine if these two instruments can be applied fairly and reliably to all learners.

I would like to invite you to allow your child to participate in the study. Should you decide to agree to your child’s participation in this research, he/she will be asked to complete the following:

1. Demographic questionnaire, which will ask information about your child’s age, etc.;
2. The Child Memory Scale; and
3. The Search and Memory Task
4. The Toulouse-Pieron Test

Participation is voluntary, and your child will not be advantaged or disadvantaged in any way for choosing to complete or not complete the questionnaire. While questions are asked about your child’s age, race and grade, no identifying information, such as his/her name or I.D number, is asked for, and as such the information will remain confidential. The completed questionnaire pack will not be seen by any person in the school at any time, and will only be processed by myself, and my supervisor. Your child’s responses will only be looked at in relation to all other responses. He/she may choose to refuse to answer any questions she/he would prefer not to, and she/he may choose to withdraw from the study at any point.

If you consent for your child to participate in the study, he/she will be asked to complete the questionnaire pack as carefully and honestly as possible. Once she/he has answered the questions, these will be collected immediately and placed in a sealed envelope. The testing

process will take a maximum of one hour and will take place at the school during school hours. There will be no risks or benefits involved in your child's participation.

Your consent for your child's participation in this study would be greatly appreciated. This research will contribute both to a larger body of knowledge on assessment in the South African context. Please do not hesitate to contact me or my research supervisor should you require further information. Should your child experience any distress after participating in the study free helpful contact numbers of counseling organizations will be provided to her/him. A feedback letter will be provided to the school and you, once I have analyzed my results. Please note that because participation is confidential I will not be able to disclose information about your child's memory scores.

Warm regards,

Riza van der Merwe

Contact details

Joseph Seabi (research supervisor) - (011) 717-8331

Riza van der Merwe (researcher) - 083 750 3762

Parental consent form

I _____ give consent for my child to partake in the study on _____.

I understand that:

- Participation in this study is voluntary.
- That my child may refuse to answer any questions she would prefer not to.
- My child can withdraw from the study at any time.
- No information that may identify my child will be included in the research report, and my child's responses will remain confidential.
- There are no risks or benefits involved in my child participation in the study

Signed: _____

Date: _____

Appendix B: Participant Assent Form

Dear Participant

My name is Riza van der Merwe and I am doing research to finish a Degree in Masters of Educational Psychology at the University of the Witwatersrand. The title of my research is “Application of the Child Memory Scale, the Search and Memory Task, and the Toulouse-Pieron Test to primary school learners”. I want to find out if these two tests really measure memory effectively. If you decide to participate in this research, you will be asked to complete Demographic questionnaire, which will ask information about your age, etc.; The Child Memory Scale, the Search and Memory Task, and the Toulouse-Pieron Test. This will take about one half hour to complete and testing will take place at your school during school hours.

As soon as the questionnaires are completed, they will be put into a sealed envelope, no-one will know who filled out what form.

I would like to invite you to participate in this study. It is your choice if you want to participate in this study, but there is no benefit or risk involved should you want to participate. If you don't want to, it's okay. If you choose to participate, any answer that you give will only be seen by my supervisor and me. In fact, you don't write your name on the questionnaire, so, we won't know who filled it out. Should you decide to participate, but there are questions that you don't want to answer, you don't have to answer them. Or if you first decide to participate, but later change your mind (at any time), you don't have to participate, either. Even if you are already filling out the form, but you decide to stop, that's okay, too.

If you have any other questions, you can contact me or my supervisor.

Warm regards,

Riza van der Merwe

Contact details

Joseph Seabi (research supervisor) - (011) 717-8331

Riza van der Merwe (researcher) - 083 750 3762

Assent form

I _____ agree to participate in the study on

_____.

I understand that:

- It is my choice to participate in this study or not.
- That I don't have to answer any question that I don't want to answer.
- I can withdraw from the study at any time.
- No information that may identify me will be included in the research report, and my answers will remain private.
- There are no risks or benefits involved for me when I participate in the study.

Signed: _____

Date: _____

Appendix C: Child Questionnaire

Code _____

Date _____

Level A _____

Questionnaire for children



Your answers are **CONFIDENTIAL** .
They will **NOT** be seen by your parents, carers or teachers.

Everyone's views are very important to us so please try to complete this questionnaire **without talking**.

If you need someone to help you **put your hand up** and a monitor will come over to you.

RANCH team, 3.15, BMS Building, Queen Mary, University of London E1 4NS. Tel. 020 7882 7649

Who are you?



1. Are you a **boy** or a **girl**?

1

Boy

2

Girl

2. What **languages** are **spoken** in your **home**?

*tick **ALL** languages that apply*

1

English

2 isiZulu

3 isiXhosa

4 Sesotho

5

Setswana

6 Hindi

7 Tamil

8 Swazi

9

Afrikaans

10 isiNdebele

11 Tshivenda

12 Xitsonga

13

Sepedi

14 SiSwati

15 Other(s) (please write)



3. What is the **main language spoken** in your **home**?

*tick **ONE** language*

1

English

2 isiZulu

3 isiXhosa

4 Sesotho

5

Setswana

6 Hindi

7 Tamil

8 Swazi

9

Afrikaans

10 isiNdebele

11 Tshivenda

12 Xitsonga

13

Sepedi

14 SiSwati

15 Other(s) (please write)



4. Has the school **offered** to give you **free school meals**?

1

Yes

2

No

3

Don't know

That's it...well done!



**Thank you for completing this
questionnaire**

Appendix D: Child Memory Scale



Story 1



Please do not turn the page until we ask you to

Story 1

Please answer the questions with as much information as you can remember from the **first** story that you heard.

1. What pieces of information did Lisa and Melissa give the police when they came?

2. What happened to the burglars and the stolen money in the end?

3. What three things did the men do to escape from the scene of the crime?

4. Who was in the brown car?



Story 2



Please do not turn the page until we ask you to

Story 2

Please answer the questions with as much information as you can remember from the **second** story that you heard.

1. Who was in the boat and what had they been doing?

2. What two events led to the boat beginning to sink?

3. When did this story take place? (Give the time of the day, the day of the week and the month of the year)

4. How did Jessica know how to save the men?



Story 1



Please do not turn the page until we ask you to

Story 1

Please answer these questions by ticking yes or no as I read them out to you. If you get left behind put your hand up and a monitor will help you catch up.

1. Yes No

2. Yes No

3. Yes No

4. Yes No

5. Yes No

6. Yes No

7. Yes No

8. Yes No

9. Yes No

10. Yes No

11. Yes No

12. Yes No

13. Yes No

14. Yes No

15. Yes No



Story 2



Please do not turn the page until we ask you to

Story 2

Please answer these questions by ticking yes or no as I read them out to you. If you get left behind put your hand up and a monitor will help you catch up.

1. Yes No

2. Yes No

3. Yes No

4. Yes No

5. Yes No

6. Yes No

7. Yes No

8. Yes No

9. Yes No

10. Yes No

11. Yes No

12. Yes No

13. Yes No

14. Yes No

15. Yes No

The letter detective

P z h
i
e



practice A

OWKJHDDYTUREZZMXDSFARNTEHJEMRAEZDBLZZNVXFDLPQHDSAKLAKMJRYKJ

SEUTGMBPLZKJVNGHTNHVBSARZEIUKUXVBZGMASHLHKBNHGMUYIUOPIPLVBXGZ

practice B

YJBMGKLPUIAEWDSOZCDFAYFGCRBVCDZSERGTBMNLKWSROXACZZDSFHRRAT

MKLPOIKMJNBGZCHF SRWCXZSFRTYEBNJMBNKVLZBSJKLPIOEAXBSNGFYRUJ

1

CXFSRQEWVZXHUYBMNMLKPWERCPVBDHGAFRXCERMQLBZNXHDYFSZWLPHGFD

XDSWBTYMCNCHZDAEQWJSCHKIYTZFDRSTCXMBNCDHFTEQRFXTSBPACWDWDMC

2

CBVHFUHOIHKLBMVZCXSWGQVQVJBKMKVKGKLPKYIUQDVOBVXBZSWGQWYEUVK

IGDSFGAHQQQLAQDMXPBGTYUQLKQPQBAFETURGIABQZQBHGDTETRQQLWQZOQ

3

TYCMBVJXZSRFGVBHMKLDBZFTYUTRFBPMKZSDIYURDAXRZXMLKAPLKRTFDN

SMGPXMZLAKDPRJQNZLKMXPBHYJTUALSMOZKAPAZPLIUJRJEOWPL

4

UZYAURGTBDUYPBNGKTUYZEMHAKKWUQZTECMBVNKKTVTMQNXBTUVIKGKLTPE

BFAAEEJGBVCZAQNKKMPKUYFUEECFBVKUHEGFVGMNKUYIERENBGIKKDNV

5

MJUYRTIOLKJEDXCVNJUZIOPLLPOIKGUHYTRECXBQHYTUIIKKNLPHCVJUIO

SVJKYMKJNBVLPOIJKUIURHDXCAVHTRHDAXBNMJFIKLPPKJNBVGCXDHWTUI

6

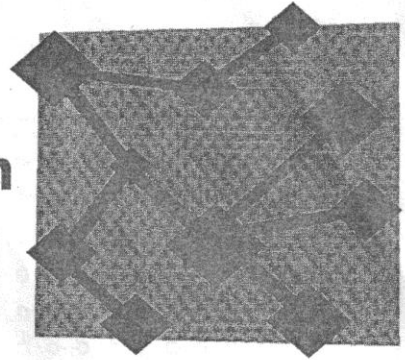
ZXCDFGBTSRPMNBXJIULPPMNBTSEFDXZCSBDYFGBNJXTSEFDXSZDFGBMHJLPPJ

BNMZGFTSBVNSUIZXNNGMMPOJIIGNBJSTEXFSSXZESCBNGIPZXMIWKGSTIB

XCVNFUBIOPLLPOILJUIBTRECVMJUBRKIOLLJEDBNICTUIJSLNLPHCVJUJO

RJDXCBVHTRJDBXBQMJUILLPPBVJGBMLJNBVLPOIJAUUIULJNBVCCXDIWTHUI

Spiky square symbol search



The spiky square symbols shown above are **target symbols**, symbols that need to be found. They are hiding in rows which are full of different spiky square symbols.

Your job is to search the rows and find the target symbols that look **exactly** like the two at the top of this box. Please search each row in turn from **left to right**.

Every time you find one of the target symbols draw a line through it like this

Then continue searching for more target symbols until you are asked to stop.

If you make a mistake put a cross through it like this

Then put a line through the correct spiky square.

Practice.

Pick up your pen and look along the rows at the bottom of this page for the target symbols.



Please do not turn the page until we ask you to



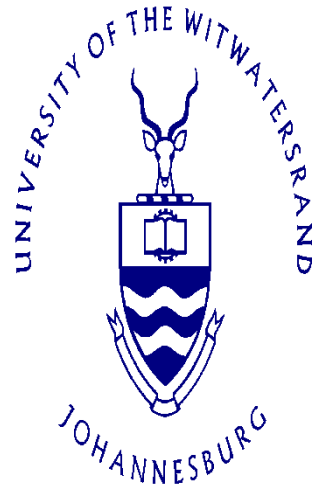
o b q b d r o o b q m o o b d r s b m r m o s b q r o o b b q r o r m o b q o m
d r o o m b d r b o q o o o m b d r q m q b m b o q o o o b o d r o q m b o b q
m o o b q r m u o o o b m o o b q o o o q o o q r o r m b b s b m r o q r o o b o
q o m b o r q b o m s b q r o o r b o o o r o b m b o q r o m b q b o o b m o o
o b q o o r o o b q m b o b o o r o m r m r d b q o b o o r q b o o r m o o o m
o o b q r o b o q b o o b m o b s b q r b o o r m b s q b m o r d b q o r m o m b
r s b o o m r o o o r o b q s b m o b b m b q o m o r q o o m b o o r q b q b
r m b s b q b m s o o m r o m b q d o r d q b o o q s b o m o q r o r o o b o o
b q o m r o o b q m r b q r s b o m o r b o o s b o m b s q b o o m b o o d r o o
b o r q r o o o q r b o o b m o o o q r o o o o b m b s q b m o b d b q o r m o m b
b q o o o m o o b r o m q o u m s b o r q r o o b r o o s b u m q m b d r b o o o q
o o m s b q r o o o o b o r m r o b b o s r o o o m r q o q b m o b d m u o o o
r d q m b o m b o r o o s b r o q b b s b m o u r q b o o o o o b o o q b o m q r o m
b m o q r o q o b r o r m o o s o r b b m o q b b o o o m o b o r q o o b o q m
r q o o o m o o o b r m q r u m s o o b q r o b o o o o m q m r o o b o o s b u o q
m o r b q o b o q o o o o m o m s o m r o q o b u s b r q r o m r b o q o s b
q r o b o o r o o m o r o q b b q m o b q o m o b m b o q o o o m o b o m r o
o o b r o o m s q b b r m o o o q m o o b q s m o q q r o b o o q o o m b q s b m o b
s b r o m b o q m o o o q b o o m o q q m q b m o o b r o r o o b o q r o o b u s b q m b o r q
m b s b q o m r b o o o m b b q o r o o o b q b o o b m r q o o m r o q r o o b
q o m r o b q b o m r q b o o b o o o r o o m b b q o r m o q r o o b m b s
o b q o o r o b q m r o o b o o s b m r o m r d b q b o o r o o r m o o r m o o o m
b o o q o b o q o o b o m o r o o r o m b o o q o o o s b u m b o q o r m q b m r
b q o o o m o o r u b m q o b o m s r o b q r o o r m b s m q m o o o o r o o q
m b r o q r o o q o r b o m b m s o b m r d e q o o o s b o r q o o m b b q s b o
q r o r o o o o m b r o q o o b q m s b q b m o s b m b o o r o r q r o o o m o b
o b o r s b q m o q b o m o b r o o q m b o b q s m r q b o o b s b m o r o o q o b
s o r b m r q q m o b r q o r b o o q o r o m q b o o b b m b q s o o m o o b
m b o b q r b o q s b r o m o b u m r d e q o o b s r o o q o o m o o b q o o r

Appendix G: Ethical Clearance Certificate for the Study

Wits School of Education

Ggg

Protocol: 2008ECE94



27 St Andrews Road, Parktown, Johannesburg, 2193 • Private Bag 3, Wits 2050, South Africa
Tel: +27 11 717-3007 • Fax: +27 11 717-3009 • E-mail: enquiries@educ.wits.ac.za • Website: www.wits.ac.za

14 August 2008

Dr. Paul Goldschagg
WSOE
L264

Dear Dr. Goldschagg

Application for Ethics Clearance

I have a pleasure in advising you that the Ethics Committee in Education of the Faculty of Humanities, acting on behalf of the Senate has agreed to approve your application for ethics clearance submitted for your proposal entitled:

Aircraft Noise and Children's Cognition and Health: a Before and After Study

Recommendation:

Ethics clearance is granted

Yours sincerely

Matsie Mabeta
Wits School of Education