THE APPROPRIATENESS OF THE RAVEN'S COLOURED PROGRESSIVE MATRICES AND ITS EXISTING NORMATIVE DATA FOR ZULU SPEAKERS

A thesis submitted in partial fulfilment of the requirements of the degree of Master of Arts (Clinical Psychology) in the School of Psychology in the University of KwaZulu-Natal, Pietermartizburg.

Tarryn Kihn

December 2005

Durban, KwaZulu-Natal

ABSTRACT

Increasing focus is being placed on fair assessment practices in South Africa. Most psychological tests used in this country were designed and developed in other countries. Research has shown that when using these tests on a different target population, it is important to determine the suitability of the test for that population. This involves investigating reliability, validity and possible sources of bias. It is equally important to have local normative data with which scores can be compared and interpreted. Research in this area has been limited. The purpose of this study was to undertake such research by evaluating a widely used test, the Raven's Coloured Progressive Matrices (RCPM), and its existing norms, for Zulu children in South Africa. The study also aimed to provide raw data, which could be used to assist with the compilation of local norms.

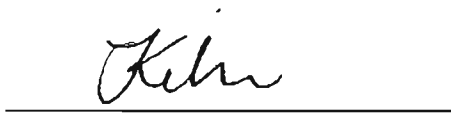
The test was administered to a convenience sample of 522 primary school children between the ages of 5 and 12 years old. The sample was divided into two groups to represent two different geographical locations: rural and urban. The sample consisted of 284 rural children and 237 urban children. There were 263 males and 259 females in the sample.

Findings indicated good test reliability for this sample. Test validity however could not be confirmed as possible sources of bias at the content and item levels were shown. Age had a significant effect on performance, but level of education appeared to be the strongest predictor of performance on the RCPM for this sample. Significant gender differences were found. There was a tendency for males to outperform females on the test. Location had a significant effect on scores, where urban children performed better than rural children. Comparison of Zulu sample scores with those produced by the British standardisation sample, revealed a large discrepancy in performance between the two different cultural groups. The Zulu sample mean scores were significantly lower than the British mean scores. These findings highlight the need for locally developed normative data. The validity of the test for the sample requires further investigation. Results indicate that test adaptation may be required for Zulu children.

DECLARATION

I, Tarryn Kihn, declare that this dissertation is my own original work. All other sources of reference have been acknowledged.

This dissertation has not been submitted previously by me for a degree at this or any other university.



Tarryn Kihn

14 December 2005

ACKNOWLEDGEMENTS

Many people have assisted me during the process of completing this piece of research. I would like to take this opportunity to acknowledge the role they played and show my appreciation for their support:

The Department of Education for their consent to conduct this study. The children and teachers of the schools involved, for their participation and assistance.

Dr. John Raven, who arranged with Jopie Van Rooyen test distributors to loan the RCPM test booklets required for conducting the research. To Jopie Van Rooyen test distributors, for their generous donation of 522 RCPM answer sheets. This study would not have been possible without these contributions.

Nicola Taylor, for her advice and help with some of the data analysis.

Vicki Nott and Jacqueline Pienaar, for their expertise on data analysis.

Bongi Qwabe for her assistance with translation of consent forms and instructions. Sibusiso Ndebele for his help with back translation.

Doug Mansfield, who supervised the beginning stages of this thesis, for his enthusiasm and encouragement.

My supervisor, Carol Mitchell, for so much: her reassurance during difficult periods of this research, her valuable experience in the field of psychological assessment, her guidance and critique. Most of all her passion, which has been a constant source of inspiration. Thank you for your care and support.

To my parents and extended family for always believing in me and helping me achieve my dreams.

And finally to my husband, Grant, for all his love and support, and the sacrifices, which made this all possible.

TABLE OF CONTENTS

CHAPTER ONE

INTRODUCTION

1.1 CROSS-CULTURAL ASSESSMENT IN SOUTH AFRICA

Since the dismantling of Apartheid and the transformation towards a democratic society in South Africa, the practice of fair assessment for all South Africans has received increased attention. The new Constitution has been particularly stringent on issues of discrimination. The Employment Equity Act No. 55 of 1998 actually outlines conditions for the use of psychological assessment measures. Specifically, psychological tests must show empirical evidence for reliability and validity and should not be biased against those being assessed (Foxcroft, Roodt & Abrahams, 2001). The Professional Board for Psychology, under the Health Professions Council of South Africa, which governs psychological practice in South Africa, sets out detailed guidelines for fair assessment procedures. The Ethical Code of Professional Conduct places the responsibility of using appropriate assessment methods on the psychological practitioner.

Despite this legislation, many assessment measures continue to be used without research establishing their applicability in a South African context. It is granted that the multicultural nature of South Africa presents unique challenges to the test user. Obstacles in cross-cultural assessment include (1) language and translation (Brislin, 1980; Reddy, Knowles & Reddy, 1995); (2) content, method and item bias (Poortinga, 1989; Poortinga, 1995; Poortinga & Van de Vijver, 1987); (3) degree of 'test-wiseness' (Nell, 1999); and a lack of appropriate local normative data (Geisinger, 1994; Helms-Lorenz & Van de Vijver, 1995). The development of new culturally appropriate tests is considered a costly initiative and should perhaps be reserved for instances where internationally recognised and accepted instruments cannot be adapted for multicultural use. Adaptation of existing tests, which have a large amount of supporting research, represents a better proposition for addressing the need for appropriate measures.

The Raven's Coloured Progressive Matrices (RCPM) is a measure of nonverbal intelligence. It was designed in the United Kingdom (UK) and has over fifty years of research behind it. It has

been reported to be a culture-fair tool, which closely approximates a "pure" aptitude test. It is a relatively cost-effective test, which according to J.C. Raven, requires very little verbal instruction (Raven, Raven & Court, 1998). Due to these favourable qualities, it is a test which is widely used in South Africa. Little research has however been conducted to establish its suitability for the various cultural groups in South Africa. Furthermore, test users have to interpret the test scores using British normative data, as the development of local norms has been limited.

It is the aim of this study to explore the appropriateness of the RCPM and its existing norms for a particular cultural group in South Africa. Zulu-speaking school children were selected as a sample for the study and were tested using the RCPM. Analysis of the raw data focused on normative comparison with the British normative data and an evaluation of the psychometric properties of the test. Of particular interest was the investigation of cultural bias and thus the validity of the RCPM for a Zulu population. Providing raw data for the compilation of local norms was considered equally critical. This study is likely to make a significant contribution to cross-cultural assessment in South Africa, an area of research which has been scarce.

Chapter two reviews the literature in the following areas: cross-cultural assessment, theories of intelligence and cognitive development; and a comprehensive review of the RCPM, its design, use, standardisation and research. Chapter three outlines the methodology used in the study. This includes the rationale and detail of sampling and administration procedures and methods of data analysis. Chapter four presents the results of the analysis and is followed by a discussion of these results in chapter five. Chapter five also considers the implications of the findings in light of the literature reviewed. Limitations of the research are explored and recommendations for further research are suggested here.

CHAPTER TWO

LITERATURE REVIEW

This chapter will provide an overview and discussion of the relevant literature and research in an attempt to contextualize the aims, hypotheses and results of this study.

2.1 CROSS-CULTURAL PSYCHOLOGY

2.1.1 A Definition

Cross-cultural psychology is concerned with

> ...the study of similarities and differences in individual psychological functioning in various cultural and ethnic groups; of the relationships between psychological variables and sociocultural, ecological, and biological variables; and of current changes in these variables (Berry, Poortinga, Dasen, & Segall, 1992, p. 2).

Research in the area of cross-cultural psychology reveals an inherent tension between seeking to understand human diversity, while also searching for universals (Berry, 1997; Berry et al., 1992). In focusing on cultural variation, cross-cultural psychologists are often involved with largely comparative studies between different cultures on specific psychological attributes. This research can be considered an example of such studies.

2.1.2 Orientations

Within the field of Cross-cultural psychology, there are three orientations which shape the way concepts and testing are used: relativism, absolutism and universalism.

Relativism is based on the assumption that behaviour needs to be understood in terms of culture and from within a culture. It avoids comparative studies and any differences are interpreted as

P0544 10

UNIVERSITY OF NATAL LIBRARY

PIETERMARITZBURG

Donation accepted for stock. Please input as an order and forward for accessioning and receipting.

Supplier DONATN

Vote PS4C

SAPSE number 20

Provenance

Special bookplate required?

Any other instructions

Initials JV.

qualitative variances in behaviour, which can be attributed to cultural variables. Relativists recommend the use of locally developed assessment procedures (Berry et al., 1992).

Absolutism occupies a contrasting position, where the importance of culture as a mediating or moderating variable is not acknowledged. It is assumed that psychological phenomena are experienced and manifested in the same way across cultures. Assessment tools are used across cultures, despite their culture of origin, and differences are understood in quantitative terms, e.g. less or more intelligent (Berry et al., 1992).

Universalism seems to lie between the aforementioned orientations. Universalists believe that there are basic psychological processes which are common to all people but recognise that these are likely to be experienced and manifested in qualitatively different ways, due to the impact of culture. Comparisons, when made, are evaluated in light of this assumption. Assessment procedures may be adapted to enhance suitability (Berry et al., 1992). The current study approaches the understanding and assessment of human behaviour from a universalist position.

2.1.3 Culture

Culture is therefore considered a core concept within cross-cultural psychology.
It is difficult to find a definition which captures the complexity of the concept of culture. In a review of many definitions, Kroeber and Kluckhohn (1952, as cited in Berry et al., 1992) suggest the following:

> Culture consists of patterns, explicit and implicit, of and for behaviour acquired and transmitted by symbols, constituting the distinctive achievements of human groups, including their embodiments in artifacts; the essential core of culture consists of traditional (i.e., historically derived and selected) ideas and especially their attached values; cultural systems may on the one hand be considered as products of action, on the other as conditioning elements of further action (p.166).

It is understood that the explicit nature of culture is readily available through observation of the everyday customs, by any observer. The principles which lie behind such customs and organize such consistent patterns of behaviour, are however more implicit and therefore imperceptible.

Understanding behaviour thus becomes complicated, not only by underlying biological and psychological processes, but by implicit cultural influences, and the dynamic interaction between all three forces. Within this interaction, culture can play a mediating or moderating role. As a mediating variable, culture could help explain the relationship between a psychological variable and a performance variable. Where culture controls or influences the strength or direction of the relationship between these two variables, it is considered a moderator variable (Baron & Kenny, 1986).

A central tenet of most cross-cultural research is that culture is one of the most critical contributing factors in differences of individual behaviour. This underpinning differentiates cross-cultural psychology from mainstream psychology, which has been both culture-bound and culture-blind.

In recognising the role culture plays in shaping behaviour, it follows that the way behaviour is studied and interpreted is also influenced by the culture of the person studying it. This is inevitable and awareness of cultural reference points is a basic requirement in cross-cultural research (Lonner & Adamopoulos, 1997). A related tendency is using one's own group's standards as the level of measurement, known as ethnocentrism (Berry et al., 1992). In order to reduce this bias, there has been a move away from research only being carried out by those living in "Western, economically and politically powerful nations" to researchers who belong to cultural groups, which co-exist in multicultural societies (Berry, 1997, p.xi).

2.1.4 The South African context

The multicultural* nature of South Africa makes it an ideal context in which to pursue cross-cultural research. The diversity of cultures within its society however, renders such research a costly and time consuming task. These obstacles have hindered research in this field to date. The consequences are particularly evident and significant in the area of psychological assessment.

* The author adopts the definition used by Berry et al (1992), which describes a multicultural society as one where cultural diversity is valued, as indicated by government policy and population attitudes.

2.2 PSYCHOLOGICAL ASSESSMENT

2.2.1 General Issues

There are many different types of psychological assessment, which employ different means of assessing various aspects of behaviour. This study is concerned with an example of a psychometric test in the measurement of cognitive functioning and the discussion of assessment will therefore be limited to this domain.

Ideally assessment should consist of a multidimensional approach to understanding behaviour. Assessment tools or measures can provide important additional information, but need to be considered in light of information gathered from other sources, for example, a clinical interview (Shuttleworth-Jordan, 1996). Psychological assessment, as based on the psychometric approach to testing, strives for objective measures. "Psychological tests are nothing more than careful observations of actual performance under standard conditions" (Russel & Cronbach, 1976, in Owen, 1998, p.13). It is important to note however, that assessment results reflect only estimates or approximations of the construct being measured. Assessment procedures are open to measurement errors, which can occur in the development of a new measure, through its use and administration, to the interpretation of its results (Foxcroft, 2001; Owen, 1998).

2.2.2 Assessment Measures

Assessment measures are directly related to their context of origin. Design of a measure is a complex process which consists of sometimes very culture specific influences, such as the construct to be measured, items selected and the preferred format of the measure.
A good psychological measure should have the following properties: "...justification for... its uses... over existing measures;...description of the sample used;... an indication of the populations, for whom the measure would be appropriate; descriptive statistics; [acceptable] reliability; [and] validity statistics" (Loewenthal, 2001, p.2).

2.2.2.1 Establishing Reliability

The reliability of a test refers to its consistency and stability, with which it measures the construct it intends to measure (Wolfaardt, 2001; Loewenthal, 2001). It can be considered "the extent to which the outcome of a test remains unaffected by irrelevant variations in the conditions and procedures of testing" (The British Psychological Society Steering Committee on Test Standards, 1992, in Loewenthal, 2001, p.5). As such, reliability can be seen to give a good indication of how accurate a test score may be. Reliability coefficients of 0.7 and above are generally used to indicate a reliable measure. Nunnally (1978, as cited in Finchilescu, 2005) recommends a much higher reliability coefficient of 0.90 when using the test results for diagnostic purposes. A slightly lower coefficient of 0.6 is acceptable if the measurement scale has less than ten items, if validity statistics are good and if the scale is appropriate on theoretical or practical grounds (Loewenthal, 2001).

Factors affecting Reliability

There are a number of factors which influence the reliability of a scale:

1. *The number of items in a scale.* The more items there are in the test, the higher the reliability of the test is likely to be.
2. *Speed versus Power tests.* In timed tests, the items are fairly similar in level of difficulty, compared to power tests where there are no time limits, but items become increasingly more difficult.
3. *Nature of sample.* A heterogeneous sample which is representative of the population under study, is likely to produce a wider range of scores, and thus gives a better estimate of reliability than a smaller, homogenous group.
4. *Range of individual differences in the sample.*
5. *Different levels of abilities which exist in the sample.* Ability levels may be indicated by age, gender or educational level, depending on the construct being measured
6. *Extraneous variables.* Procedures in test administration are likely to either enhance (if standardized) or compromise (if not standardized) the reliability of a test.

(Finchilescu, 2005; Wolfaardt, 2001)

Reliability is a necessary condition for the validity of a test, but does not ensure validity. If a test however, is unreliable, it is not likely to be valid (Rosnow & Rosenthal, 1996).

2.2.2.2 Establishing Validity

Once reliability of an instrument is established, it is necessary to ascertain whether it measures what it claims to measure, and how well it does this. This is known as validity. The three categories of validity are: content, criterion-related and construct validity. Content validity refers to whether the items in the test cover the domain of the construct being measured. Criterion-related validity is present if the performance on the test can be positively correlated with performance on a similar measure (concurrent validity) or can predict future performance (predictive validity). Construct validity is the most difficult to measure as it involves determining whether the test actually measures the intended theoretical construct. This requires investigating whether the operationalisation (how the construct is translated into observable behaviour) does indeed represent that construct and is the best possible way of defining it. (Finchilescu, 2005; Geisinger, 1994; Loewenthal, 2001; Rosnow & Rosenthal, 1996; Wolfaardt, 2001). Consider, for example, the difficulty in measuring the theoretical construct of intelligence. An exploration of this process is covered in 2.2.6.2.

Factor analytic techniques are commonly used to evaluate construct validity (Poortinga, 1995). Factor analysis involves examining the underlying factors, dimensions or traits, which the items in the test load highly on (Hammond, 1995). Geisinger (1994) notes that when applying factor-analytic comparisons across cultures, large representative samples should be drawn and variables like age, gender and educational level (when using tests of cognitive performance) should be controlled. Some findings suggest that the same dimensions or factors exist across cultures but that the loadings of items on factors may differ (Geisinger, 1994). Irvine (1979, as cited in Poortinga, 1995) however, has questioned whether the same factor-analytic structure of intelligence, in particular, could be the same across cultures. His argument has been revived more recently by those who advocate a relativistic approach to assessment (Miller, 1995, in Poortinga, 1995).

Content and criterion validity are difficult to measure cross-culturally because they are related to factors which are naturally culture dependent. For example, identifying criteria of school success is likely to be different not only across cultures but also across different schools within the same culture (Geisinger, 1994).

Information on validity informs the test user as to what the performance scores on that test actually mean. Interpretation of scores is made with the aid of normative data. Typically, after designing an instrument, it is tested on a representative population sample in order to compile standardized performance scores. These standardized scores are known as normative data. Comparisons across ages and groups can be made so that the individual can be placed in relation to the population distribution of scores, which are assumed to be normally distributed (Geisinger, 1994; Hammond, 1995; Wolfaardt, 2001). The importance of using normative data which is representative of the population, from which the testee is drawn, is highlighted by Van de Vijver and Poortinga (1991, as cited in Geisinger, 1994) in their observation that "Scores that are numerically identical can have a psychologically different meaning across cultures" (p.310).

2.2.3 Adaptation of Measures

The International Test Commission (Version 2000, in Foxcroft, Roodt & Abrahams, 2001) outlines fair and ethical practice in assessment. It places the responsibility of appropriate test use on the test user and highlights the importance of considering the impact of broader socio-political and cultural factors on test results, their interpretation and application. A test score must always be interpreted in light of the context from which it was obtained (Claassen, 1997, in Bedell, Van Eeden & Van Staden, 1999).

The lack of standardized psychological tests and normative data is a critical problem in South Africa. Test performance is influenced by several variables. These variables, which include cultural and socio-economic background, educational level and individual ability, may prevent normative data, collected from one population, from being used as the yardstick for another. Although many researchers have acknowledged and highlighted this deficiency (Anderson,

Poortinga & Van de Vijver, 1987; Scheuneman, 1979). Cultural equivalence is indicated by the absence of bias and refers to whether the construct in question has a similar meaning across the different cultural groups under study.

One such method is test adaptation. The International Test Commission has published guidelines for adapting assessment measures (Foxcroft et al., 2001). In reviewing the literature on test adaptations, Helms-Lorenz and Van de Vijver (1995) found that many researchers focus on improving administration procedures, like clarifying and simplifying test instructions and providing more practice examples. While there is also a lot of work done on improving the validity of tests for specific groups, they found that there are not many studies which concentrate on improving the appropriateness of tests for multicultural settings (Helms-Lorenz & Van de Vijver, 1995). Test adaptation requires careful analysis of the test on all levels, from design to use. By investigating and eliminating possible sources of bias, a test's suitability for a new target population is increased.

2.2.4 Assessment Bias

Assessing cultural equivalence in testing relies on identifying and eliminating bias which may be present in the design of the test and / or in the use of the test. The success of any measurement tool depends on its design, how it is used and the abilities of the tester in interpreting its results (Owen, 1998). While the test developer is responsible for design, reporting on reliability and validity and providing standardized normative data, it is the psychologist's responsibility to use the test appropriately. If there are changes in the way the test is used, it is necessary to reestablish reliability and validity. Changes in test use would include when the test is altered and when a test is applied to a new population. It must be investigated whether the test continues to measure the same construct and whether consistency is maintained. In considering whether a test is appropriate for an individual or group thus requires investigating possible sources of bias (Geisinger, 1994).

Taylor and Radford (1986, as cited in Bedell et al., 1999) argue that "bias can never be eliminated entirely" (p.2). Poortinga (1995) believes that "the pursuit of unbiased instruments is

a fiction" (p.144). The testing context, subject, tester, the test and test items all present possible sources of error. Bias is defined as a constant error rather than a chance error (Anastasi, 1990). Once identified, these systematic errors may be minimized and controlled. Cultural bias is evident when the test results from one culture have a different meaning to those interpreted in another culture (Pootinga, 1995).

2.2.4.1 Types of Bias

Research into test bias focuses on three sources of bias in particular: construct bias, method bias and item bias (Helm-Lorenz & Van de Vijver, 1995). Method bias refers to factors which influence test performance, such as familiarity with testing procedures. Bias at the item level of the test is known as item bias. This term has more recently been replaced by the term differential item functioning (DIF), where item analysis focuses on whether the items function differently for different groups. DIF indicates that "...individuals having the same ability, but from different groups, do not have the same probability of getting the item right" (Kanjee, 2001, p.97). DIF focuses on the analysis of inter-item correlations and item difficulty values (Poortinga, 1995). This is linked quite closely to the reliability of a test, since inter-item and item-total correlations are used to indicate internal consistency of the assessment measure. Both reliability and validity of a test are therefore automatically improved, once DIF is identified and those items are removed (Kanjee, 2001). This type of item analysis is often used in cross-cultural research and is advocated for tests being used in multicultural settings, in particular (Helms-Lorenz & Van de Vijver, 1995; Van de Vijver & Leung, 1997). Poortinga (1995) warns however, that item bias analysis is not likely to detect an underlying bias or variance which is shared by all items.

When using a test across different groups, one of the most important questions to ask is whether the same construct is being measured across those groups. If it is not, the test is biased in terms of the construct. This is similar to investigating whether a construct is valid for a particular group. Owen (1998) cites the following criteria as indicators of the absence of construct bias: "similar test reliabilities in the two cultural groups; similar rank orders of item difficulty values; similar item discrimination values (and) similar factor structures" (p.62).

The probability of all three types of bias being present in an assessment, is increased when test results are used to generalize and predict performance in domains other than those specifically measured by the test (Helms-Lorenz & Van de Vijver, 1995). For this reason, Helms-Lorenz and Van de Vijver (1995) consider construct bias to be the most problematic in aptitude tests. (The distinction between aptitude and achievement tests is discussed in 2.2.6.2 under Intelligence tests)

2.2.4.2 Cultural Loading

The term cultural loading refers to a specific cultural context, which is found in the culture of the original test developer, in the instrument or in administration of the test. Properties in the instrument itself, coupled with characteristics of the sample used, will contribute to the cultural loading on the test (Helms-Lorenz & Van de Vijver, 1995).

Van de Vijver and Poortinga (1992, as cited in Helms-Lorenz & Van de Vijver, 1995) identify five aspects of assessment, which can contribute to cultural loading:

1. *Tester.* When the tester is from a different culture to that of the test-takers.
2. *Test-takers.* Factors related to the test-takers as a group, otherwise known as subject-related factors. These will be explored in more detail next.
3. *Interaction between tester and test-takers.* If there are communication problems between the tester and test-takers, the administration procedure could be compromised, which in turn affects the validity of the test.
4. *Response protocols.* This includes time limits as well as unfamiliar test formats.
5. *Test items.* The suitability of items for different cultures has been reviewed under 2.2.4.1 "Types of bias".

Many subject-related factors impact on test performance. Comparisons of results must be interpreted with caution when the following factors apply:

1. *Proficiency of language used in testing.* This relates more to tests which are highly dependent on verbal ability, but can also be applied to the instructions given in any testing procedure.
2. *Cultural Norms.* There are often implicit references to dominant cultural norms in test items.
3. *Test-wiseness.* This refers to test-taking skills, like speed, concentration and motivation. Nell (1999) claims that "...in psychological assessment, the most fundamental difference between westernised subjects and those from non-western cultures is their degree of test-wiseness" (p.129).
4. *Acculturation strategy.* Acculturation can have either positive or negative effects on performance. This is a special area of interest within cross-cultural psychology and within multicultural assessment, in particular. It is discussed further in the following section.

(Helms-Lorenz & Van de Vijver, 1995)

These factors can contribute to the three types of bias described in 2.2.4.1, although the probability is not equal across the different forms of bias. Subject-related factors like verbal skills and test-wiseness are likely to affect all test items in a consistent manner. It is these cross-cultural differences, which Poortinga (1995), claims will not be identified in item bias. These factors fall under method bias and are most likely to invalidate intergroup comparisons (Helms-Lorenz & Van de Vijver, 1995). Poortinga & Van de Vijver (1987) note that "it can remain quite unclear whether an observed intercultural difference is valid, or due to bias, even if the data do pass one of the usual tests for bias" (p.278).

It seems that bias analysis is not likely to fully resolve the issue of cross-cultural comparability. Poortinga and Van de Vijver (1987) suggest that an explanation of cross-cultural differences is needed once bias has been identified. This is facilitated by including variables which may account for the bias effects into the design of the research project (Poortinga, 1995) Cultural

bias is thus seen as "not some form of measurement error, but a systematic component of variance that needs to be explained" (Poortinga, 1995, p.145). This approach changes the goal from comparing different cultural groups on a psychological attribute to rather collecting information on, and understanding the nature of cross-cultural differences. Moreover, if cultural equivalence cannot be demonstrated, it can be regarded as an indication that improved operationalisations are required (Poortinga, 1989).

2.2.5 Special Issues in Multicultural Assessment

As mentioned previously, test adaptation in multicultural settings has received little attention and has no set guidelines to follow. The multicultural setting presents unique challenges to the researcher, which makes test adaptation a difficult and complicated task. Possibly the largest challenge in conducting this kind of research in a multicultural context, is the process if acculturation (Helms-Lorenz & Van de Vijver, 1995; Shuttleworth-Jordan, 1996).

2.2.5.1 Acculturation

Changes in cultures are natural and inevitable processes which happen over time. Change may be in response to numerous internal and external factors. Acculturation is an example of change which occurs when one culture comes into contact with another.

> Acculturation comprehends those phenomena which result when groups of individuals having different cultures come into continuous first-hand contact, with subsequent changes in the original culture patterns of either or both groups... (Berry et al., 1992, p.271).

There is the possibility of mutual influence between the two cultures, but in reality where one culture is considered to be more dominant than the other, the influence is not equal. Acculturation creates change which is unpredictable and highly variable. It affects individuals and whole cultural systems. Psychological changes are expected at an individual level, while political and economic changes are expected at a group level (Berry et al., 1992). Berry (1992)

observes that "virtually any behaviour studied by psychology is a candidate for a shift during acculturation" (p.281).

Acculturation is thus a difficult process to study. Of relevance to this particular study, are the following points:

1. The dominant culture exerts more influence over the younger generation of the non-dominant culture because 1) the older generation is more entrenched in their culture and therefore more resistant; 2) children undergo formal schooling, which is usually organized by the dominant culture. Other indicators would include urbanization, access to media and the level of formal schooling. Furthermore, these are all likely to be interrelated. Acculturation is therefore viewed as a "multilinear phenomenon" (Berry et al., 1992, p. 278).

2. Research shows that cognitive performance is often positively correlated with acculturation (Berry et al., 1992). It is difficult to assess whether there have indeed been new cognitive operations developed or whether the individual has just become more learned in the dominant culture's test-taking behaviour. Cognitive change as a consequence of acculturation is therefore often difficult to specify.

3. Acculturative Stress: The process of acculturation is often experienced by individuals as stressful and can have a negative effect on their physical, psychological and social well-being

For these reasons, Helms-Lorenz and Van de Vijver (1995), believe that acculturation strategies exert powerful influences on an individual's adjustment and on their expectation for the future. These influences have direct bearing on the educational setting.

2.2.5.2 Differential Norms

One method of compensating for test bias and realizing fairness in a multicultural setting is to use differential norms. This involves adjusting scores so that different cut-off points are used for different cultural groups. This is a highly controversial practice, which has wider social and political implications. Affirmative action and equal opportunities policies are examples of the application of differential norms. Helms-Lorenz and Van de Vijver (1995) however, argue that this method should be used when bias in a test is detected. This is especially relevant with cognitive tests, which are used diagnostically and have the power to control access to resources which have long-term social and economic consequences for the individual test-taker.

2.2.6 Cognitive Assessment

The area of cognitive assessment, in particular, compared to other areas of assessment, has generated huge amounts of research and interest because of the value modern society tends to place on cognitive ability. Cognitive tests are used qualitatively and diagnostically in education. They identify which children pass and which children fail. They help to diagnose learning disorders and mental retardation, and motivate placement decisions like special schooling (Helms-Lorenz & Van de Vijver, 1995). Cognitive test results are often used as selection criteria for entrance to tertiary educational institutions and success in job applications (van Eeden & de Beer, 2001). As such, cognitive tests wield a certain amount of power over an individual's access to resources. In clinical settings cognitive tests aid diagnosis of organic and non-organic brain dysfunction and have implications for medical treatment of the individual. In South Africa, cognitive assessment contributes to clinical practitioners' recommendations for disability grants funded by the state social welfare system.

Before exploring cognitive assessment in more detail, it is necessary to consider the definition of cognition and the prevailing theories of its development. Nutrition is considered an important context variable in cognitive assessment and its impact on cognitive development is reviewed. This is followed by a discussion of theories in intelligence and how they are applied to intellectual assessment.

2.2.6.1 Cognitive Development

"Cognition refers to the inner processes and products of the mind that lead to "knowing". It includes all mental activity – remembering, symbolising, categorising, problem solving, creating, fantasising and even dreaming." (Berk, 2000, p.221). The development of cognitive functioning has been studied extensively, with many, varying theories suggested. This discussion will focus on Piaget's theory because it is one of the most widely recognised perspectives and because it is relevant to the RCPM, which is being used in the current study. Vygotsky's theory has received increasing credit and is considered here because of the importance placed on the child's cultural context.

Piaget's Cognitive-Developmental Stage Theory

According to Piaget, the accumulation of knowledge is a process where the child is an active participant in its cognitive development. He thus saw intelligence or mental growth as a constant effort to gain and improve on knowledge. The child must draw upon current psychological structures, or schemes, to interpret his world. In interacting with the environment, the child's psychological structure is adapted or modified. Adaptation consists of two complementary processes of assimilation and accommodation. Assimilation involves using current schemes to interpret the world, while accommodation involves adjusting or creating new schemes in light of environmental influences. These schemes change with age according to Piaget's four stages of development:

1) *sensorimotor stage* (birth to 2 years):
   The infant uses complex sensory and motor schemes to gain more control over his environment. Behaviour becomes more intentional and object permanence is achieved.

2) *preoperational stage* (2 to 7 years):
This stage is marked by changes in the capacity for representation and symbolic activity. The child is still not yet capable of operations and reasoning ability is limited.

3) *concrete operational* (7 to 11 years):
This stage is characterised by more logical and flexible reasoning. Operations like conservation are mastered as well as a variety of spatial operations.

4) *formal operational* (from adolescence to adulthood)
The child is now able to reason abstractly and use problem-solving strategies and hypothetico-deductive reasoning.

(Berk, 2000)

Piaget considered these stages of development to be universal and the sequence as being genetically determined and therefore fixed (Berk, 2000). He acknowledged that the "social milieu" within which the child is raised affects the rate of progress through the stages, but gave little attention to identifying and explaining which factors might play a role (Thomas, 1985). More recent research has found that many of the operations Piaget identified within each stage, actually emerge earlier than he suggested and are mastered earlier than he expected (Berk, 2000). Cross-cultural research has supported Piaget's concept of stagewise development, where the same sequences of behaviours have been reported across cultures. There were however large individual differences observed, and the stages did not necessarily occur in order of chronological age (Berry et al., 1992; Jahoda & Krewer, 1997). Other findings have led experts within the area of child development to seriously question the classification of cognitive development into discrete stages (Bjorklund, 1995 & Flavell, 1992, in Berk, 2000). Furthermore, research has suggested that specific experiences (like formal schooling) may be critical to the mastery of Piagetian tasks, especially those in the concrete and formal operational stages. This challenges Piaget's view that operational thinking is not dependent on particular

kinds of experience. Culture and training seem to be important conditions of, and influences in cognitive development (Berk, 2000).

Vygotsky's Sociocultural Theory

While Piaget gave little attention to the role of culture, Vygotsky viewed it as an integral factor in the structure of cognitive development. Unlike Piaget, he did not believe that intellectual patterns of thinking are primarily genetically determined, but rather they are products of the interaction between the child and activities valued by his society. "It follows that practical thinking will predominate in societies that are characterised by practical manipulation of objects, and more 'abstract' forms of 'theoretical' activity in technological societies will induce more abstract, theoretical thinking" (Luria, 1976, cited in Thomas, 1985, p. xiv). Vygotsky was influenced by Marxist theory, in particular historical materialism, which states that historical changes in society create change in human nature (Vygotsky, 1978).

Vygotsky initiated the term "zone of proximal development" to refer to the range in a child's abilities between independent and mediated performance on tasks. Independent performance reveals actual ability, while mediated performance reflects the potential development under guidance (Owen, 1998; Vygotsky 1978). Vygotsky's theory of child development thus views cognitive ability as dynamic, not static, and dependent on specific learning experiences. It also attempts to provide a framework which recognises potential ability.

Vygotsky regarded language as an essential foundation for the development of higher cognitive processes. Language represents a "culturally produced sign system", which mediates individual development (Vygotsky, 1978, p.7). He believed that the "language climate" the child is raised in, at school and at home, affects the level of conceptual thinking reached. By identifying language as a tool of culture, Vygotsky was the first psychologist to show how culture becomes internalised (Vygotsky, 1978).

Nutrition

The role of nutrition in optimal cognitive development has long been recognised and the effects of malnutrition well documented (Berk, 2000; Lezak, 1995; Lynn, 1989; Whaley et al., 2003). Prenatal malnutrition can seriously impair the development of the central nervous system of the foetus. Early brain development is more susceptible to malnutrition than at later ages of childhood. Research has shown that "the poorer the mother's diet, the greater the loss in brain weight" (Berk, 2000, p.104). Later, during middle childhood, these children are likely to perform poorly on intelligence tests. Severely malnourished children (those with height or weight-for-height which is less than -2 Z-scores) show significantly impaired reasoning and perceptual-spatial functioning, poor school performance and attentional deficits. Mild to moderate malnutrition also produces significant impairments in intellectual functioning (Whaley et al., 2003). Many studies have shown a direct relationship between malnutrition and low IQ scores, although malnourishment is likely to be associated with other factors like poverty and low parental intelligence, which may also be responsible for this correlation (Lynn, 1989). Identical twin studies where the twins had different birth weights, have provided greater control for confounding variables found in genetic and environmental differences. Low birth weight, due to inadequate prenatal nutrition, affected later intellectual performance. Interestingly the significant difference in lower IQ scores for the lighter twins, has been found on performance IQ measures, with little difference observed in verbal IQ scores (Willerman & Churchill, 1967 & Hendrichsen, Skinhoj & Andersen, 1986, in Lynn, 1989).

Theories around the role of nutrition have changed over the years, where the focus in the 1970s on protein deficiency has now been adjusted to a focus on poor diet quality and a lack of micronutrients (Whaley et al., 2003). A Kenyan study measured the impact of dietary supplementation on grade one school children's cognitive performance. Comparisons were made across three groups of supplementation (animal source foods (ASF), milk and high energy) and a control group, on tests of cognitive ability, over a three year period. Their findings show that the group receiving ASF outperformed the other groups on the RCPM and arithmetic tasks, with the performance on the RCPM demonstrating the most significant gains. Supplementation did not have an effect on verbal performance. The results suggest that meat provides a better source of

micronutrients than the other supplements. Context variables such as SES showed that children from higher SES performed better, regardless of supplementation group, than the other children on all cognitive tasks. It may be assumed that families of high SES are likely to have more food and food which is of a good quality (Whaley et al., 2003). Improved nutrition and increased non-verbal IQ performance has been observed on a global level over the past few decades (Lynn, 1989; Raven et al., 1998).

The trend known as the Flynn effect has been noticed in studies conducted in westernised nations, and shows that children's intelligence has been increasing. The increase has been greater for visuo-spatial abilities than verbal abilities (Lynn, 1989; Raven et al., 1998). In fact Raven et al. (1998) report that increases in performance on the Raven Coloured Progressive Matrices has been approximately 2.5 IQ points per decade. Lynn (1989) suggests that improvements in nutrition in wealthier Western nations are responsible for the increase in intellectual abilities in children. Her theory relates to the findings from twin studies that performance IQ tended to be more susceptible to poor nutrition than verbal IQ. Improvement in prenatal and early childhood nutrition is likely to result in improvement in performance IQ scores (Lynn, 1989). Raven et al. (1998) agree with a "nutrition, hygiene" hypothesis, noting that the largest gains in IQ scores have been in the low socio-economic status groups (Bouvier, 1969, in Raven et al., 1998, p.G18). Nutritional status must therefore be considered an important context variable in cognitive performance (Poortinga & Van de Vijver, 1987).

2.2.6.2 Intellectual Assessment

Theories of Intelligence

There is much controversy around the concept of intelligence and its definition and many theories have been posited in an attempt to reach a more accurate explanation of intelligence. It is beyond the scope of the present study to explore this debate and discussion will focus on theories which are relevant to the RCPM design.

Initial investigations in the area of intelligence testing date back to the late 1800's, where Sir Francis Galton discovered that human traits tended to be normally distributed (Owen, 1998). Since then there have been many theories on what intelligence is and how it could be measured. There have also been numerous tests developed, which claim to measure an intelligence quotient (IQ). The majority of these instruments have been designed from a psychometric approach. This approach is based on factor analysis which was invented by a British psychologist named Charles Spearman (1863-1945). His theory of intelligence involves a general factor "g" for general intelligence and specific factors "s". Positive correlations between activities indicates the g factor. Psychometric intelligence tests therefore focus on items which correlate with g. They are standardized, quantitative and focus on final total score rather than strategies used to tackle test items (Owen, 1998).

Eductive ability is one of the two components which Spearman identified as forming the underlying common or general factor of intelligence or g. The second component being reproductive ability, involves recalling mostly verbal material based on a culture store of knowledge and ideas. Horn (1994, as cited in Raven et al., 1998) found that both components, (often referred to as fluid and crystallized intelligence respectively) are distinct from an early age but have different genetic origins and are influenced by different environmental factors. They do however interact, where the process of perception, of organizing and making meaning of new visual material, relies on a store of already acquired concepts. Similarly learning and recalling verbal information is dependent on making meaning out of a confused array of verbal constructs (Horn, 1994, in Raven et al., 1998). The models proposed by Spearman, Cattel and Horn respectively are examples of the more enduring theories within the field. There is however, no unanimous definition of intelligence and because there are so many theories, many authors in the field, like Lezak (1997) believe that the term IQ has become meaningless. Meaning can possibly be preserved and isolated in context, where the use of a particular assessment measure provides results which reflect certain aspects of intelligence as identified within a particular culture (van Eeden & de Beer, 2001).

Intelligence Tests

Ability testing can be divided into aptitude and achievement tests. The difference lies in the measuring of learnt abilities versus abilities which are less dependent on specific learning experiences. School achievement tests measure skills and knowledge which are acquired through education. This has been referred to as crystallized intelligence (Cattell & Butcher, 1968, in Helms-Lorenz & Van de Vijver, 1995). In contrast aptitude tests are supposed to rely more on fluid intelligence, which is less dependent on previous learning experiences. Intelligence tests are examples of aptitude tests, although research has shown that they more often involve both crystallized and fluid intelligence, in the degree to which they use some verbal ability and academic knowledge (Cattell & Butcher, 1968, in Helms-Lorenz & Van de Vijver, 1995). The Raven's Coloured Progressive Matrices (RCPM) has been identified in the literature, as an intelligence test, which closely approximates a "pure" aptitude test because it uses "...simple stimulus material that is not acquired in school" (Helms-Lorenz & Van de Vijver, 1995, p.160).

Non-verbal tests such as the RCPM are frequently used in assessment in South Africa, due to the language difficulties, which arise in many clinical settings. Wober (1969) however, advises against assuming a test is culture fair because it is non-verbal. Traditional cross-cultural tests have been based on the assumption that nonverbal content is a more culture-fair approach to intellectual ability assessment (Owen, 1998). Implied within this statement is the idea that nonverbal content can measure the same functions as verbal content and according to Anastasi (1990), this is highly questionable. Moreover, Olson (1986, as cited in Owen, 1998) states that if the above suggestion is true, it proves that nonverbal content is equally culturally biased. In fact, some studies show that nonverbal tests in comparison to verbal tests may actually be more culturally loaded, especially if the concepts used are unfamiliar (Anastasi, 1990; Crawford-Nutt, 1976).

Since most intelligence tests have been developed in Western societies, they tend to measure abilities which are held as important and necessary by Western societies. Kendell and colleagues (1988, as cited in Bedell et al., 1999) warn of the danger in using these tests as universal measures of intelligence. After all " ...culture dictates what will be learned, by whom, and at

what age" (Ferguson, 1956, in Bedell et al., 1999). Intelligent behaviour in an African context would have been defined by the person's ability to hunt, provide for his family and participate in the oral tradition of his ancestors (Zindi, 1995, in Owen, 1998). These are not skills which are valued in Western society and some authors argue that it may be inappropriate to use standard Western measures of intelligence in Africa (Bedell et al., 1999). Helms (1992) suggests that performance on intelligence tests is merely a reflection of how well the subject has mastered "White culture". Indeed Raven et al. (1998) found that in pre-literate societies like those in Africa, exposure to two-dimensional or symbolic representations of shapes and lines is not customary for many people. The RCPM then becomes problematic when used with people who do not have familiarity with these concepts, as the matrices problems depend on the ability to perceive and think about abstract designs. Furthermore, it appears that personal motivation to engage with abstractions varies, and thus values (as indicated by motivation) tend to affect RCPM scores more than J.C. Raven thought (Raven et al., 1998). These findings highlight the role of experience and culture in cognitive ability.

2.3 RAVEN'S COLOURED PROGRESSIVE MATRICES (RCPM)

2.3.1 Test Design and Use

The RCPM measures, amongst other things, intellectual maturity. It is considered a sensitive tool in detecting right hemisphere brain dysfunction, as over a third of the test items measure visuospatial competency (Lezak, 1995). It is designed for use with young children and the elderly, and can be used with people who do not speak nor understand English, as it is a nonverbal measure. Although instructions for the test are outlined in the manual, Raven et al. (1998) claim that instructions are not a necessity, as the problems to be solved are self-evident. This has however been disputed by various authors using the test in an African context (Wober, 1969). The test is not timed, and on average takes between thirty to forty minutes to complete (Raven et al., 1998).

There are two forms of the test; a book and a board form. Both are brightly coloured to make the test more attractive for children. The board form of the test consists of movable pieces, which enable the subject to physically manipulate and experiment with alternative options. There are no reported differences in results using the two different forms. It has however been observed that intellectually impaired subjects have difficulty in understanding the book form of the test and respond better to the board form. Generally, it seems that when confronted with representations in a novel way, these subjects have difficulty interpreting and using symbols. In the board form of the test, orientation has also presented problems for these subjects (Raven et al., 1998).

Both forms of the test consist of a series of patterns or matrices, from which there is a piece missing. The subject must select one of six options presented at the bottom of the diagram to make the pattern complete. These 36 designs or matrices are divided into three categories or sets. The problems in each set are arranged so that the difficulty increases across each set and so that each problem provides learning for completing successive problems. The sequence of items has been reviewed on several occasions, using item analysis, to ensure rank ordering of difficulty (Raven et al., 1998).

Subjects must use reasoning by analogy to infer the correct responses (Raven, 1965). This type of problem solving ability relies on eductive ability, which may be defined as "making meaning out of confusion; developing new insights; going beyond the given to perceive that which is not immediately obvious; forming (largely non-verbal) constructs which facilitate the handling of complex problems..." (Raven et al., 1998, p.G4). The variables of the relationship are not readily recognizable and so the relationship must be "educed" in order to identify the variables. In other words the whole must be seen first to be able to understand the particular. Problems in Set Ab (the second set) in particular, require the subject to see the whole in order to work out the missing part. Raven et al. (1998) consider mental development in childhood as "more like salmon leaps in the stream of life than equally spaced rungs of a ladder" (p.CPM1). Reasoning by analogy seems to be a critical stage in intellectual maturation and Raven et al. (1998) maintain it is usually reached by 11 years of age. This stage is particularly vulnerable to any kind of organic dysfunction. If there has been impairment or deterioration in intellectual capacity, the CPM is a useful tool in assessing clarity of thought and the level of impairment or deterioration.

(The test cannot indicate stability of already acquired skills and knowledge nor the capacity for new acquisition.) This makes it a suitable test for use with people who have sustained brain damage or are of subnormal intelligence (Raven et al., 1998). The Raven's Standard Progressive Matrices (RSPM) has been developed to assess children older than 11 years and adults.

It must be noted that the RCPM was not designed as a test of general intelligence but rather of nonverbal intellectual ability. Raven et al. (1998) recommend the combined use of the Chricton Vocabulary scales (CVS) for a full measure of intellectual functioning. Correlations between the CPM and CVS indicate that a combination of the two provides a good measure of general intelligence (Raven et al., 1998).

2.3.2 Standardisation

The RCPM has been restandardised twice since its inception in 1947; in 1979 and more recently in 1982 (Raven et al., 1998). Of special mention is an observation that across the standardizations there has been acceleration in development (as discussed under "nutrition" in 2.2.6.1). The most recent standardization in 1982 was conducted in Dumfries, Scotland, on 598 children. It included children from special schools. Younger children and those who experienced difficulties were tested individually, while the others were tested by group.

The RCPM has also been normed in European countries (including Germany, Switzerland, Belgium, Holland, Portugal, Spain, France, Italy, Finland, Sweden, Poland, Austria,Yugoslavia and Hungary), English speaking countries (including Australia, New Zealand, Canada, United States of America (USA), United Kingdom (UK)), South and Central American countries (including Guatemala, Peru, Uruguay) and countries in the Middle and Far East (including India, China, Taiwan, Korea, Hong Kong, Singapore, Malaysia, Greece, Egypt, Israel, Turkey, Lebanon, Iran).

Norms for Australia, West Germany, Slovakia, Switzerland and Spain are all similar to those of the UK sample. Higher norms were found in East Germany and the Netherlands, while Sao Paulo in Brazil produced lower norms. A study carried out in 1993 in an isolated mountain area

of Peru, has produced the lowest norms ever reported. The USA studies reported higher norms for high socioeconomic districts and lower norms for more rural areas (Raven et al., 1998). These variances across samples highlight the importance of collecting local normative data. Indeed after reviewing four very large studies conducted in the USA, Sigmon (1983) concluded that "...local norms would be best for specific populations" (p.486).

There is currently little normative data available for the elderly. Efforts to collect normative data in South Africa are presently underway (J.Van Rooyen, personal communication, November, 2004). The results of the current study will be added to this database.

2.3.3 Reliability

The RCPM has a reported overall split-half reliability estimate of .90. There seem to be varying results around the reliability for different age groups. Carlson and Jensen (1981) reported very low reliability for the age group 5.5 to 6.5 year olds (Cronbach alpha = .57; split-half = .65). This follows earlier findings by Harris (1959, as cited in Raven et al., 1998) of split-half reliability of .466 for the same age group. Most studies however, tend to report values around .80 (Court & Raven, 1995). Sample size, which varied across these studies, is likely to play a role in different reports.

With specific reference to non-English speaking samples, there has been a high reliability (Cronbach's alpha = 0.94) reported across several studies (Li, Hu, Chen, Jin & Jeng, 1988; Teng, Li, Chen & Jin, 1991; Miao & Huang, 1990, cited in Raven et al., 1998). In Kenya, Costenbader and Ngari (2001) calculated a Cronbach coefficient alpha of 0.87 on their standardization sample of 1370 children, which was composed of 46 different tribal groups. In South Africa, Kendall (1976) found good reliability (KR20 = .87) for rural Africans and Bass (2000) reported Cronbach's alpha of .88 on a small sample (n=379) of Xhosa pupils.

Test-retest reliability correlations for the Raven's Progressive Matrices fall in the range of 0.7 to 0.9 (Eichorn, 1975; Llabre, 1984, in Lezak, 1995) "...even when retesting involves three administrations six and twelve months apart." (Lezak, 1982c, in Lezak, 1995, p.614) Test-retest

reliability studies also support stability across cultures with reliability coefficients ranging between .81 to .92. Length of retest interval plays a role in variance where estimates tended to drop as the interval period increased (Costenbader & Ngari, 2001; Raven et al., 1998).

Carlson and Jensen (1981) found the RCPM to be consistent across three different ethnic groups (Anglo, Hispanic, and Black). Internal consistency, as measured by item analysis, in a study conducted by Green & Kluever (1991, in Court & Raven, 1995) was reported at .89 on an American sample. Raven (1983, cited in Court & Raven, 1995) reported item difficulty correlations of .97 on a very large sample of 3250 children between the ages of 6 and 16 years in the United Kingdom (UK) and .98-1.00 in an American sample of 30,000 children of the same age range. Removing items which have a very low or high p value is possible without compromising internal consistency, although it is unclear how this would affect validity (Carlson & Jensen, 1981).

2.3.4 Validity

Analysis of the internal structure of the test using factor analysis has been investigated by several authors. Wiedl and Carlson (1976) identified three factors, through the use of principal component analysis, which accounted for 36% of the total variance of their sample of 180 children. Factor 1 was interpreted as *concrete and abstract reasoning* and is associated with items which are ranked as the most difficult in the test. (Carpenter, Just and Shell (1990) confirm that those problems requiring abstract rules are the more difficult problems). The second factor is *continuous and discrete pattern completion* and consists of items which are considered the least difficult, while Factor 3 contains items of medium difficulty and can be defined as *pattern completion through closure* (Wiedl & Carlson, 1976). Although this finding indicates a rather heterogeneous test structure, the combined loadings of discrete with continuous pattern completion, and concrete with abstract reasoning items, suggest that tasks depend less on discriminatory processes than Raven implied.

In a later study, using a larger sample (n=783) and two different methods of factor analysis, Carlson and Jensen (1980), corroborated Wiedl and Carlson's findings. The factor defined as

Concrete and Abstract Reasoning was modified to *Closure and Abstract Reasoning by Analogy,* after using tetrachoric correlations, since this method of analysis seemed to reduce the impact of item difficulty on factor structure. Schmidtke and Schaller (1980) produced a similar factor structure using principal component analysis with varimax rotation. Regardless of the coefficient used, these studies suggest that the Raven's Coloured Progressive Matrices test items load on three specific factors. Furthermore, these factors vary in the way they draw on perceptual and conceptual processing abilities. Schmidtke and Schaller (1980) note that "the resultant factors may reflect item characteristics as well as processing requirements of the tasks" (p.1246).

Simultaneous processing has been frequently implicated in the solution of CPM problems, with loadings in the range of .75 to .85 (Kirby & Das, 1978; Das, Kirby & Jarman, 1975; Carlson & Wiedl, 1976; Raven et al., 1998). Simultaneous processing has been shown to be highly related to spatial ability, and less related to reasoning and memory. It involves the integration of many stimuli simultaneously in order to form a gestalt or grasp relationships between two or more objects (Kirby & Das, 1978). This form of processing has also been indicated in performance on Piagetian tasks, specifically those measuring concrete operational thought. Conservation concepts and class inclusion, in particular, have been related to the perceptual items in the CPM (Carlson & Wiedl, 1976; Raven et al., 1998). Carlson (1973) divided subjects into two different groups according to problem solving strategies used, as identified by the Inhelder and Piaget's standard multiple classification Matrix test. Subjects were then administered the RCPM. Their performance revealed that different parts of the RCPM seem to rely on different processes, namely Set A and Ab use graphic processes, while Set B relies on operational methods.

2.3.4.1 Cross-Cultural Studies

Klippel (1975) on the other hand claims that different abilities are required in tackling Piagetian tasks compared to those needed to solve problems in standard intelligence testing. Her study assessed three different ethnic groups in New Zealand (Maori, Samoan and Pakeha (European descent)) on Piagetian tasks, the RCPM, Stanford-Binet test (S-B) and Peabody Picture Vocabulary Test (PPVT). CPM correlations of .41 with the S-B and .26 with the PPVT are reported. Interaction effects between gender and ethnicity made it difficult to state definitively

whether the CPM is "culture-fair". Van de Rijt (1990, as cited in Helms-Lorenz & Van de Vijver, 1995) explored culture loadings on the RCPM, Wechsler Intelligence Scale for Children, revised (WISC-R) and a verbal IQ test. The RCPM showed the lowest loadings of the three tests, but overall the correlation between effect size and cultural loading was significant. Differences in performance between natives and migrants in the Netherlands increased as cultural loadings increased. In a study comparing performance on the Wechsler Intelligence Scale for Children (WISC) and the CPM, Martin and Wiechers (1954) found correlations of .91, .84 and .83 between full scale, verbal and performance IQ's respectively. Correlations between Information (.47) are the lowest with Block Design the highest at .74. Lower correlations (r = .68) were reported when comparing the RCPM with the Spanish version of the Wechsler Adult Intelligence Scale (WAIS) (Fletcher, Todd & Satz, 1975). An unpublished South African study assessing mentally impaired adult Coloured subjects with the RCPM and the South African Wechsler Adult Intelligence Scale (SAWAIS), demonstrated a statistically significant positive relationship between performance on the two tests. The highest correlation appears to be with the Digit symbol subtest (.57) closely followed by the Block Design subtest (.54). Information indicated a negative relationship (-.08) (Maree, 1994).

Emerling (1990) found that the RCPM may be a biased predictor for Mexican American children. His sample is small (n=115) and seems to contradict other findings by Hoffman (1983, in Raven et al., 1998), although Emerling had equated his groups for socioeconomic status (SES), which Hoffman had neglected to do. Once groups are equated for SES, Emerling (1990) found they are comparable in terms of nonverbal intelligence and achievement. This supports work done by White and Panunto (1978) comparing non-verbal performance of first and second language children. They found that when SES is controlled, and a familiar setting is used in testing, the second language children's RCPM performance is comparable to English speaking children. The RCPM was included in a battery of tests used to assess predictive validity with the Classification Test Battery (CTB) for South African mineworkers. The intercorrelation between the CTB and RCPM was calculated at 0.61.

The predictive effect of gender on performance of the CPM appears inconclusive. There have been inconsistent findings in several studies. Some findings suggest that males outperform

females (Bass, 2000; Costenbader & Ngari, 2001; Klippel, 1975; Sigmon, 1983), while Jinabhai and colleagues (2004) found females' performance to be superior.

In two large African studies (one conducted in Kenya by Costenbader & Ngari (2001) and one in South Africa by Jinabhai et al., (2004)), the only measure used to assess the validity of the RCPM was observing whether individual responses behaved as expected for the test, that is, whether the number of correct responses decreased as test items become more difficult through the sets. These researchers have tended to rely on validity evidence from other studies on different populations, to support their rationale for using the test on their target populations. This is a concern, considering how lower mean test scores have been repeatedly obtained in African samples when compared to Euro-American test norms.

In the standardization of the RCPM in Kenya, children between 6 and 11.5 years differed substantially in their performance from the norms produced by the UK and USA. Furthermore item analysis revealed that item A7 of the test proved to be more difficult for the Kenyan sample and thus sequencing of the test items is questioned for this sample (Costenbader & Ngari, 2001). Jinabhai et al. (2004) found that rural Zulu children in South Africa performed at three to five years below children used in international norms. These results seem to be surprisingly consistent with a very old study conducted by Notcutt in 1949 with Zulu children, which showed a deficit of approximately three years in comparison to the English sample. Similar depression of scores was observed in a study investigating Xhosa speaking primary school children in South Africa (Bass, 2000).

More studies in South Africa tend to involve the use of the RSPM, and tend to report results which are consistent with those obtained using the CPM. Owen (1992) observed large mean score differences (almost three standard deviations) between black and white Grade 9 students. Black university students fared no better when compared to overseas age counterparts (Grieve & Viljoen, 2000, in Jinabhai et al., 2004). Interestingly, Crawford-Nutt (1976) found that black students performed in line with British norms, although her sample size was small (n=114). Zimbabwean students have also scored significantly lower than British students by almost two standard deviations on the RSPM (Zindi, 1994).Vass (1992) experimented with different

methods of test presentation of the RSPM in his study of Xhosa secondary school children. Results showed that the group receiving the alternate form of instruction, including visual aids and extra practice items, scored significantly higher than those receiving the standard Raven instructions. Scores overall however still fell well below those in international samples (Vass, 1992). In a study assessing the effects of Feuerstein's Mediated Learning Experience on RSPM scores, African university students showed significant improvements, although their performance still fell below that of the non-African students (Skuy et al., 2002). Vernon (1967) and Lloyd and Pidgeon (1961) in South Africa, Silley (1964) in Uganda and Jahoda (1956) in Ghana, all confirm that practice and coaching show gains in performance on the RCPM (in Wober, 1969). In light of these studies, Wober (1969) claims that a second testing of the CPM provides a more valid reflection of true performance. These results tend to suggest not only the importance of familiarity with test taking skills, but also familiarity with the visual medium of the test content. In fact, Sinha (1968, as cited in Raven et al., 1998) concluded that lower performance among his Indian subjects was the result of the test's heavy reliance on visual-spatial abilities

Following many of the findings which demonstrate the extent to which the RCPM relies on visual-spatial abilities, this discussion would be incomplete without a review of the research in the area of perceptual development and differences.

2.3.5 Perceptual Skills

Research into cross-cultural differences in visual perception has provided evidence that not only is familiarity with pictorial materials important but that direct experience with such materials is necessary.

Research in this area of cross-cultural psychology tends to be quite dated. Nevertheless, findings are relevant to more recent research with the RCPM. Serpell (1979) states that one of the problems in this area of research is that results which indicate differences in performance across different cultural groups, has been translated into differences in underlying psychological processes and structures. Cross-cultural research has shown how these differences in visual perception can actually be related to familiarity with pictorial materials and exposure to formal

- education (Ardila & Roselli, 1994; Berry, 1971; Hudson, 1962a; 1962b; Miller, 1973; Serpell, 1979)

Spatial perceptual development in children is understood as being shaped by many mediating factors. Biological determinants are influenced by nutrition, socialization processes and larger cultural and ecological factors (Berk, 2000). Berry (1971) proposes a model which emphasises the interactions between these variables. The type of perceptual skills developed depends on environmental demands on the person and their cultural adaptation to them. The role of culture within this model consists of aids like language coding, arts and crafts and socialization (Berry, 1971). The technique of socialization has been indicated in the development of certain cognitive styles (Dawson, 1967, in Miller, 1973; Witkin 1962, in Berry, 1971). It seems that children, whose parents use an authoritarian style, where obedience and conformity are valued over initiative and independent thinking and behaviour, tend to be "field-dependent" in their approach to a perceptual field. By contrast field-independence is stimulated by encouragement of individual interests and ideas. This fosters a more analytic approach in perception (Dawson, 1967, in Miller, 1973; Witkin 1962, in Berry, 1971).

In a study comparing traditional and transitional (those involved in a process of acculturation) groups within the same culture, Berry (1971) found gender and age trend differences in spatial abilities. It seems that males perform better than females where there are strong gender roles within a culture. Furthermore the usual developmental curve for performance across age is only found where those particular skills are needed to meet particular ecological demands (for example visual discrimination is highly developed for hunting). A rather flat developmental curve showed that there is little growth in perceptual skills, which are not considered as useful. As a result of his study, Berry (1971) concludes that the level of visual discrimination and spatial ability will develop according to ecological demands and not to an inherent universal developmental pattern.

More recent research on the differences observed between the genders on performance of certain cognitive abilities, like spatial and perceptual tasks, has been viewed from an evolutionary

aspect. The division of labour through human history is thought to play a role in the natural selection of abilities valued by societies through the evolution of the human species. These abilities are therefore not only socially reinforced but may also have been perpetuated through genetic evolution (Kimura, 2000).

Serpell's (1979) "specific experience" hypothesis is similar in its premises to that of Berry's (1971) idea of ecological demands. Serpell (1979) tested his hypothesis by comparing English and Zambian children across four tasks which varied from testing universal experience to more specific experience tasks. It was predicted that English children would perform better on a drawing task, while Zambian children would outperform the English children on a wire-modelling task. The results matched predictions. Serpell (1979) uses his results to argue for a narrower construct than field-dependence or field-independence to explain differences between African and western performance. If it were a valid construct, it does not explain why the English children who scored high on the drawing task should not achieve equal success on a similar pattern reproduction task in the medium of wire. Likewise, for the Zambian children, Serpell (1979) suggests that it is not representational abstraction which is difficult for them in the drawing task, it is merely the medium used.

While familiarity with materials and tasks has been found to greatly enhance performance, some authors stress the role of direct experience in interpreting representation (Hudson, 1962a; 1962b; Miller, 1973; Serpell, 1979). African children in Hudson's (1962b) study were able to identify and recognize elements in drawings, but not able to work out correctly the spatial relationships between elements. Serpell (1979) concludes that "in the non-verbal domain of pattern reproduction, the medium of testing interacts with the subject's specific prior experience" (p.379). Miller (1973) confirms that the more abstract the representation becomes, the more experience with pictures is necessary. Serpell (1972, as cited in Serpell, 1979) cautions that "To measure with non-verbal pictorial tests the abilities of children whose cultural environment does little to impart pictorial skills is just as hazardous an enterprise as testing children in a second/non-dominant language (p.3790).

Ardila and Rosselli (1994) found that socioeconomic level, defined as the "quantity and quality of environmental stimulation" a child will receive was significant but that this influence decreased as the number of years of education increased (p.111). In assessing the impact of education on visuospatial and memory abilities, Ardila, Rosselli and Rosas (1989) compared performance on a neuropsychological battery between illiterates and professionals. Gender differences were only observed in the illiterate group. This finding was explained by how many of these people functioned according to traditional gender roles. Literacy, which could be understood as specific training in cognitive abilities, has important consequences for perception, logical reasoning and remembering (Laboratory of Comparative Human Cognition, 1983, in Ardila et al., 1989). Berry (1971) and Serpell (1979) assume that cognitive processes are universal and differences are ascribed to different cultural environments and ecological demands. Ardila et al. (1989) however question the assumption of universal inborn cognitive abilities. They propose that in assessment, one is measuring abilities which are specific and highly trained. Ardila et al. (1989) believe that cognitive abilities are learned and education is thus a more significant variable than culture.

2.4 SUMMARY AND AIMS

Research is increasingly focusing on the fair use of psychological assessments in cross-cultural applications (Berry et al., 1992; Foxcroft, 2001; Claassen, 1997, in Bedell et al., 1999, Nell, 1999; Helms-Lorenz & Van de Vijver, 1995; Poortinga & Van de Vijver, 1987; Emerling, 1990; Geisinger, 1994). In South Africa, this type of research is still lagging, despite ethical and legal obligations for practitioners to use tests appropriately. The present study attempts to evaluate a test (RCPM), which is widely used in neuropsychological and cognitive assessment across South Africa, in terms of its appropriateness for a Zulu-speaking population. The study aims to compare the test sample's performance with the most commonly used British normative data. The appropriateness of the British data for interpretation of Zulu scores is questioned. In addition, the study intends to investigate the psychometric properties of the RCPM and whether it is indeed a valid test for this sample. The aims and hypotheses are outlined in more detail in

the following chapter. This is followed by a description of the methodology employed in this study.

CHAPTER THREE

METHODOLOGY

This chapter outlines the proposed aims and hypotheses of the current study. In addition, the characteristics of the research design, sample and procedure will be presented.

3.1 AIMS AND HYPOTHESES

It is the aim of this study to investigate the appropriateness of the RCPM and the existing normative data for a particular cultural group in South Africa. The researcher has conducted the research in KwaZulu Natal and has chosen Zulu children as the target population. The Zulu people represent the largest population in KwaZulu Natal (South African Statistics, 2004/05).

Peripheral aims would be to:

1) provide initial normative data on the local population that could be more valid than the existing norms,
2) show the difference between existing norms and local norms and explore possible explanations for these variations,
3) show the relationships between variables such as age, gender, education and geographic areas and their impact on test performance, and
4) explore the suitability of the test for the local population

Not only will this study aid neuropsychological and educational assessments, by providing interim local normative data, but it will also add literature to a field within which, research has been scarce.

3.1.2 Research Questions and Hypotheses

"Are existing norms for the Raven's Coloured Progressive Matrices appropriate for Zulu speakers?"

"Is the Raven's Coloured Progressive Matrices an appropriate test for Zulu-speakers?"

Hypothesis 1: There will be a significant discrepancy, taking age into account, between the performance of the study sample's subjects and the performance of the RCPM standardisation sample.

Hypothesis 2: Age will have a statistically significant effect on performance, where a positive correlation is expected

Hypothesis 3: Educational level will have a statistically significant effect on performance, where a positive correlation is expected

Hypothesis 4: Gender will not have a statistically significant effect on performance

Hypothesis 5: Geographical area will have a statistically significant effect on performance (rural < urban)

3.2 RESEARCH DESIGN

A mixed model design, which included elements of a survey, correlational and quasi-experimental design, was used to conduct this study. A method of non-random sampling was employed in the form of convenience sampling, due to limited financial and personnel resources. It is noted that this was an opportunity sample with no control for variables such as learning disorders, school failure, head injury, epilepsy or severe emotional or behavioural problems.

Children were weighed and measured for height in order to calculate a body mass index (BMI). (BMI = weight (kg)/height (m)^2.) The BMI was used as a covariate to add information relating

to levels of nutrition. Nutrition plays an important role in cognitive development and performance, as has been shown in 2.2.6.1. The BMI is noted as an indicator of nutritional status. It is used as an effective screening tool to assess underweight and overweight individuals. It is however, not a diagnostic tool. BMI for children is gender and age specific because children's body fat content changes as they grow and these changes differ for boys and girls (Centres for Disease Control and Prevention, 2005). BMI charts produced by the National Centre for Health Statistics and the National Center for Chronic Disease Prevention and Health Promotion (2000) were used to assess the BMI scores of the children in the study. (These are used in district hospital settings by dieticians in KwaZulu-Natal (J. Galliers, personal communication, September 12, 2005)).

Independent variables are age, educational level, gender and geographical area. The dependent variable is test performance.

3.3 SAMPLE

3.3.1 Description of Study Area

This study was conducted in the province of KwaZulu-Natal, which has a population of 8.4 million, of which 60% is rural. Africans comprise the largest ethnic group in the province (81.8%), with the largest proportion belonging to the Zulu nation. The number of school children in the province has been estimated at 2.7 million. (Jinabhai et al., 2004; South African Statistics, 2004/05). Pietermaritzburg is the capital city of the province, with a population of approximately 350,000 (South Africa Online Travel Guide, 2005). The sample was drawn from two different geographical locations within the areas surrounding the city. The geographical locations represent a rural and urban sample. The urban location was in the township of Imbali, which is situated within 30 kilometres (km) of the city centre. New Hanover, a small farming community situated approximately 100km from Pietermartizburg, was chosen as the rural location.

There have been difficulties in locating an accurate and agreed upon qualitative definition of what constitutes a rural, versus an urban area. There are varying criteria that range from

population density to economic activity. For the purpose of this study, it appears more relevant to focus on the differences in school quality in these two areas. A rural school is thus considered a school which is located some distance from a main city or town, where effects of acculturation would be considered minimal. School infrastructure is expected to be inferior to that in urban schools. (For example, teacher - student ratios are expected to be higher, level of teacher training lower and resources/facilities limited). Improvement on all these dimensions would be expected in an urban school environment. Research by Card and Kreuger (1996, as cited in Kaufman, Maharaj and Richter, 2001) and Loeb and Bound (1992, as cited in Kaufman et al., 2001), has explored the relationship between school achievement and school quality and found a positive correlation between these variables in developing country settings. Studies also show that access to resources and mere exposure to an urban environment, contributes to test-wiseness (Nell, 1999). Acculturation is considered a powerful influence in test performance and can be operationalised in terms of geographic area (Shuttleworth-Jordan, 1996), where those living in an urban environment are considered to have more exposure to acculturative factors like contact with the dominant culture and the media (Berry et al., 1992).

Differences in standards of housing were observed between the two locations. The children attending the rural school, came from either the informal settlement next to the school or from workers' compounds on neighbouring farms. The homes were constructed mostly with mud and thatch or tin roofs. In Imbali, a range of buildings exists, from shack dwellings to architecturally constructed homes of brick and mortar. Both locations have access to electricity and running water, although financial constraints would limit some occupants' use of electricity. It is important to consider that 49 % of people of employable age in the province earn less than R800 per month (Statistics South Africa, 2001).

Differences in the schools were also noted. The rural school buildings were in poor condition, with many broken windows. There was a shortage of desks and overcrowding in the Grade one classrooms. Teacher-to-pupil ratios ranged between 1: 25 to 1: 45. Resources like writing instruments and paper are limited. The school had a store of textbooks, although it is unclear how often pupils are taught with the aid of books or pictorial materials.

There appeared to be greater use of textbooks in the classrooms of the urban school. Writing instruments and paper were also in greater supply. The school buildings were in better condition and there were more buildings to accommodate the children, although some cramped conditions were also noted in this school. Teacher-to-pupil ratios were in a similar range to that observed in the rural school.

Both schools lacked sports facilities, but did have large open areas for play. Both schools benefit from feeding schemes, although both principals reported that this creates a stigma among the older children and so it is only generally the younger children who take advantage of the food available. The food provided, consists mostly of samp and beans, and vegetables are limited to cabbage and carrots. There is little animal protein in the meal (N. Mzila, personal communication, November, 2004).

3.3.2 Participants

522 Zulu speaking children between the ages of 5 and 12 years old were drawn from schools in each of the areas. The school used in the rural location yielded 284 participants, while the urban sample contained 237 participants. Although raw scores are analysed according to chronological age, participants were tested by grade (Grades 1 through to grade 5). This distinction was made because chronological age is not synonymous with educational level. A range of ages within each grade was observed. There were 263 males and 259 females in the sample.

Samples may be biased by absenteeism. Those who were absent during the period of testing were excluded from the sample.

The children in the rural sample came from homes where parents or carers were either unemployed, employed as farm workers on farms in the area, or receiving disability or pension grants. Parents and carers from the urban sample, tended to show a range in occupational functioning, from unemployment and government grants to teachers and civil servants.

3.4 INSTRUMENTS

3.4.1 Raven's Coloured Progressive Matrices

The RCPM test was used. The design of this instrument and its psychometric properties are discussed in Chapter two.

3.4.2 Electronic Scale

An electronic scale for body mass was used to record the weight of each participant. Weight was recorded in kilograms.

3.4.3 Tape Measure

Height was measured with a tape measure and recorded in centimetres.

3.5 PROCEDURE

3.5.1 Ethical Approval

The Department of Education was contacted for permission to conduct research in the two schools selected for the sample. After permission was granted, the principals of the schools were approached for their permission to collect data in their schools. (Appendix C contains a copy of consent from the Education Department.) Consent from parents/guardians of each child participating in the study, was obtained. (The consent form in English and Zulu can be viewed in Apendices A and B respectively.) Consent forms were distributed by teachers to pupils two weeks prior to data collection. Of the consent forms sent out, only three parents refused consent for their children to be involved.

3.5.2 Nonresponse and Absenteeism

Those children who had not returned forms by the start of the data collection, were given replacement consent forms and were tested the following day in another group. Children with consent, who were absent on the day of their grade's testing, were included in another group when they returned to school. There were five children who were absent during the entire period of testing and were therefore excluded from the study.

3.5.3 Administration

The duration of test administration varied with each grade. Grade 1 required more time with explanation of instruction procedures and completion of the test. On average an hour per group was necessary. As grades increased, time for administration and completion decreased, ranging from 40 minutes to 15 minutes.

Testing was conducted over the course of the school day (7:30am – 13:00pm). Approximately one day per grade was necessary to complete administration within that grade. Data collection was carried out over a period of one week in each school.

The RCPM test was administered by group. Each grade was divided into groups of 8, as recommended in the RCPM Manual (Raven et al., 1998). Where possible, children were spread out so that there was an empty desk between each subject. (Some classes were overcrowded and pupils shared desks. In such cases, subjects had their own desk, but were not separated by an empty desk). Pupils who were not participating in the group procedure at any one time, were occupied with work set by the teacher. The groups were carefully monitored and children often had to be reminded to do their own work. Each child received individual RCPM test booklets and answer sheets to record their responses. Each child had their own pencil. (In the larger classes, pencils are sometimes shared between two pupils for class work. This was particularly evident in the rural Grade 1 classes). There was a great deal of excitement observed in most of

the children tested, due to the research situation. The researcher was of different race and culture to the students, using new materials and requesting their help in furthering research in their country. The experience was a novel one for many of the children. This was especially evident in the rural Grade 1 classes, who were the first participants in the study, and witnessed the unpackaging of all the new books and answer sheets. (Receiving or using new goods can be considered a rare experience for many of these children).

Instructions as per Raven et al., (1998) were closely adhered to. They were translated into Zulu and read to each group. (See Appendix D for the translated instructions.) Translation from English into Zulu involved the method of back translation, as recommended by Brislin (1980), to ensure accuracy and prevent compromise of testing procedure. Preceding instructions for the test, there was an explanation of the research rationale and participants' role in the study. They were assured that the test was not for school purposes and would not affect their school performance. They were also assured that access to the results of the test would be limited to the researcher and parents, where necessary.

Assistance of the teacher in each grade was negotiated. The teachers were fully briefed on the purpose of the test, instructions for test administration and how they may deal with students' queries. This precluded the cost of a translator without compromising test instruction comprehension. Test anxiety is expected to have been minimized by 1) using the teacher (a familiar person) and 2) conducting the administration in the classroom (a familiar setting).

Most children, especially the younger ones, displayed a great deal of difficulty in understanding the answer protocol. Grade one pupils required lengthy and repeated instruction in how to record their answers. Repeated instructions for the completion of the test were also necessary. After instruction, those children who continued to display confusion, were given individual assistance by the examiner.

3.6 ANALYSIS OF DATA

Data entry and analysis was completed with the use of SPSS version 11.5. Data was analysed from two perspectives, namely normative evaluation and internal analysis.

3.6.1 Normative Evaluation

Normative evaluation involved straightforward statistical comparisons between the original normative sample from the UK, and the study sample. The existing normative data was compared with a) rural area, (b) urban area and (c) total study sample, according to age. An independent *t*-test was employed to make these comparisons, since the UK normative data is considered to be based on a sample which is different from the study sample, and thus based on a different population. This required establishing that the present study sample was normally distributed and that variance was homogenous (Durrheim, 2005). Each age group was converted to Raven's age categories to facilitate comparisons. A trend analysis was computed to measure the rate of change or increase in performance across age groups. (A linear progression in raw scores over age groups tested would be an indication of test validity.)

3.6.2 Internal Analysis

Descriptive statistics of the sample were calculated by age, gender, grade and geographical location.

Internal analysis assessed the relationship between participant attributes and performance on the RCPM, in order to observe whether the norms for Zulu-speaking children can be reported without respect to age, gender, educational level and geographical area. These attributes represent factors which are likely to interact with each other and therefore complicate any linear relationship between one attribute and performance, which may be observed. Multiple regression

analysis was used to identify the relationships between the different independent variables and to assess the significance of their predictive abilities in relation to the dependent variable.

3.6.2.1 Reliability

Internal consistency was calculated using Cronbach's coefficient alpha for the final data set. Cronbach's alpha is the most widely used reliability coefficient and is regarded as the most important index of test reliability (Cronbach, 1951 and Nunnally, 1978, cited in Loewenthal, 2001).

Test-retest reliability could not be calculated because the test was only administered on one occasion to the sample. This would have been a useful estimate of reliability and would have added to the overall weight of reliability findings.

Item analysis was conducted to observe whether the individual items correlated with each other, and with the total score on the whole scale. This reflects the cohesiveness of the scale's internal structure. It also indicates items which make a poor contribution to the scale and what the resulting alpha coefficient would be if the item was removed from the scale (Loewenthal, 2001).

3.6.2.2 Validity

Discrepancies at the level of item analysis contribute to test bias and indicate differential item functioning, where items may be inappropriate to a particular context (Van de Vijver & Leung, 1997). By inspecting item-total correlations, it is possible to assess the discriminatory power of an item. Item-total correlation coefficients of 0.2 and above are considered good indications of items which correlate well with the total scale and have discriminatory power (Foxcroft, 2001; Loewenthal, 2001). Negative item-total correlations highlight those items with very poor discriminatory power (Foxcroft, 2001).

The difficulty of items may also vary in different sample populations. As early as 1966, Irvine suggested that "where a test measures different abilities in different cultures, item ranks (of

difficulty) will differ" (cited in Wober, 1969, p.233). Currently, there is a debate over which model (one parameter vs. three parameter model) is most appropriate for the DIF analysis on the RCPM (N. Taylor, personal communication, November 24, 2005). Due to this uncertainty and limited access to resources, DIF was not performed on the raw data. The proportion of correct responses for each item for each age group was instead calculated. This will provide a superficial view of the item functioning and may make DIF speculation possible.

Principal component analysis was also performed on the items, in order to identify the underlying dimensions of the test for this sample. Although this is not a representative sample, it is a large one and therefore the use of a factor analysis technique is justified (Geisinger, 1994). This analysis isolates which items load on those dimensions or factors. It is suggested that principal component analysis may give clearer answers than factor analysis (Loewenthal, 2001). There are two different methods of rotation, which can be used in this analysis; orthogonal and oblique. The choice depends on the correlation of the underlying factors (Hammond, 1995). It is suggested that factors on an intelligence test like RCPM would be correlated and that principal component analysis with oblique rotation would therefore be indicated. Other studies investigating the factor analytic structure of the RCPM have however, used an orthogonal rotation, namely varimax technique. It was decided that two separate analyses, each using one type of rotation would be run. This would serve to show any differences found using two different types of rotation and provide results which could be easily compared to other findings.

Identifying a similar factor structure to that already identified by previous research findings, would provide evidence for construct validity (Van de Vijver & Leung, 1997). "Cattell has always been fully aware of the necessity of comparing factor structures across cultures before using the test..." (Geisinger, 1994, p.398).

3.7 ETHICAL ISSUES

It is recognised that this sample consisted of children who represent a vulnerable population. The design of the study attempted to minimise harmful effects to the children. A familiar environment and the use of familiar persons during testing would have reduced test anxiety. Test

anxiety is considered the only harmful effect of this study. Full informed consent was obtained, as detailed in 3.5.1., and the caregivers and children were free to withdraw their consent at any time during the study. Confidentiality of the schools and participants has been maintained by omitting their names from the study report. Participants' results will only be released to their respective caregivers. Participants and their caregivers were likely to benefit from the study, as it provided the potential to identify children with problems and refer them to the appropriate resources.

The results of the data analysis are presented in the following chapter.

CHAPTER FOUR

RESULTS

This chapter summarises the results of the data analysis as outlined in chapter four. Chapter five will provide a discussion of these results and implications for future research.

4.1 DESCRIPTIVE STATISTICS

Five hundred and twenty-two pupils participated in this study. There were 259 (49.6%) female pupils and 263 (50.4%) males in the total sample. The sample was subdivided into half-yearly intervals, according to Raven's age categories. Raven's CPM normative data is presented in this format. A description of the sample in terms of age and gender by age categories can be found in Table 1 in Appendix E. The sample consisted of an urban and a rural group, with the rural group comprising 285 of the total participants and the urban, slightly less, with 237 participants. Table 1 reflects the characteristics of the sample.

Table 1

Total Sample Characteristics

	Gender		Grade					Total
Age	Male	Female	1	2	3	4	5	
5	2	0	2	0	0	0	0	2
6	16	11	27	0	0	0	0	27
7	22	33	45	9	1	0	0	55
8	47	39	34	41	11	0	0	86
9	61	56	13	35	67	2	0	117
10	56	63	10	19	40	44	6	119
11	41	42	2	6	7	41	27	83
12	17	13	0	3	4	13	10	30
13	1	2	0	1	1	1	0	3
Total	263	259	133	114	131	101	43	522

Table 2

Means and standard deviations of age and grade by geographical location

		Mean	n	SD
Age	Rural	9.29	285	1.61
	Urban	8.56	237	1.42
Total		8.96	522	1.57
Grade	Rural	2.69	285	1.44
	Urban	2.56	237	1.04
Total		2.63	522	1.28

Table 2 contains the means and standard deviations for age and grade by geographical location. The participants ranged in age between 5 and 13 years, with a mean age of 8.9 (standard deviation = 1.6). The highest number of participants fell into Grade One, where a total of 133 children were tested. The age within this grade had a large range; between 5 and 10.5 years. A breakdown of the sample by grade is presented in Table 3.

Table 3

Description of sample by grade

				Gender	
Grade	n	Mean Age	Age range	Male	Female
1	133	7.2	5-10.5 yrs	75	58
2	114	8.6	6.5-12.5 yrs	59	55
3	131	9.1	7-12.5 yrs	62	69
4	101	10.3	9-13 yrs	49	52
5	43	10.8	9.5-12 yrs	18	25

4.2 NORMATIVE EVALUATION

In this section the descriptive statistics (including the frequency distribution and the mean, standard deviations, range and maximum and minimum scores) for the sample are presented.

4.2.1 Mean, Standard Deviation and Range of Scores

The mean score, standard deviation, range, minimum and maximum scores for the sample are presented in Table 4.

Table 4

Mean, Standard Deviation, Range, Minimum and Maximum Scores

Mean	SD	Range	Minimum Score	Maximum Score
16.69	5.73	29	4	33

The mean score for the 522 participants in this sample was 16.69, with a standard deviation of 5.73. The range was between 4 and 33. A score of 36 on this test represents a perfect score.

4.2.2 The Distribution of the Scores

The frequency distribution of the scores for the sample is presented in table (Table 5) and graph (Figure 1) format.

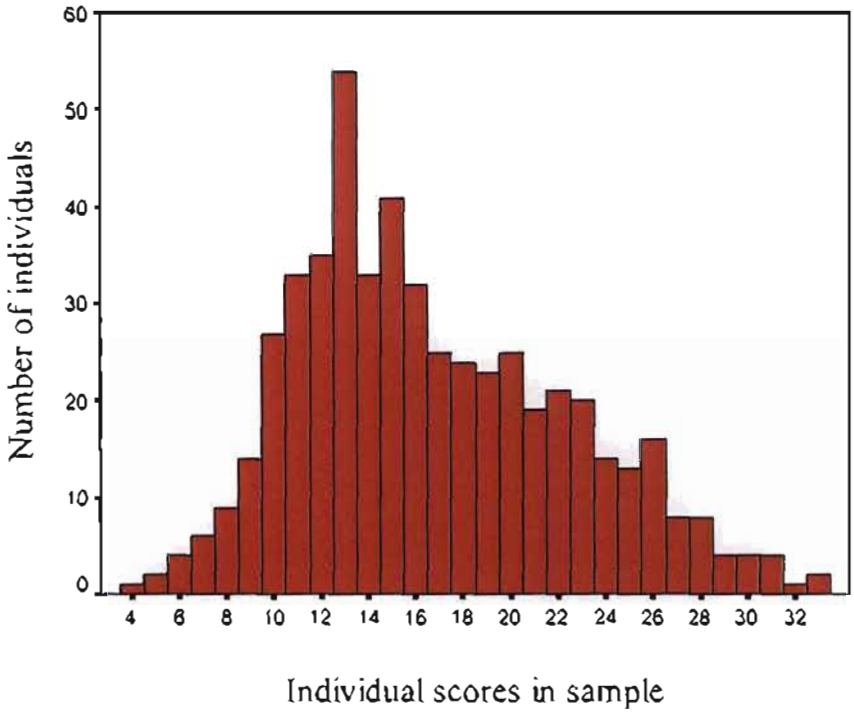
Table 5

Frequency Distribution of Scores on the RCPM

Score	*f*	Cumulative Frequency	Cumulative Percentage
0 -4	1	1	0.2
5 - 8	21	22	4.2
9 - 12	109	131	25.1
13 - 16	160	291	55.7
17 - 20	97	388	74.3
21 - 24	74	462	88.5
25 - 28	45	507	97.1
29 – 32	13	520	99.6
33 - 36	2	522	100.0

Figure 1

Graph of Frequency Distribution of Scores on the RCPM



It is evident from the graph that there is a positive skew in terms of the distribution of scores on the RCPM for this sample. The scores are therefore not normally distributed. There is a sharp increase in the frequency of scores from the lowest score of 4 to the maximum frequency at a score of 13. This is followed by another peak in the frequency of score 15. There appears to be a steady decline in the frequencies of scores thereafter, with only two performances of 33 on the test. It is important to note that 33 was the highest score achieved on the test, with no results producing a perfect score of 36. Moreover, 55.7 % of the sample scored below the mean.

4.2.3 Hypothesis 1

This South African sample was compared with the most recent Raven British standardisation sample (1982). The South African sample was organised according to Raven age categories to

facilitate comparative analysis. The British sample only covers ages 5.5 to 11.5. The age categories of 5, 12 and 12.5 were therefore omitted from the South African sample, for the purposes of comparison. Table 6 displays the means and standard deviations of individual scores for both samples.

Table 6

Group Statistics

Population	n	Mean	SD
British sample	13	23.85	6.24
South African sample	13	16.35	1.92

On inspection of the table, it is apparent that there is a large discrepancy in the differences of the mean scores of the two samples. The British sample has a much higher mean score of 23.85, than the South African sample mean score of 16.35. The South African mean score has dropped from 16.69 of total sample due to exclusion of three age categories. The British and South African samples were compared across 13 age categories, ranging from 5.5 years of age to 11.5 years of age, at half-yearly intervals. The results of the t-test (presented in Table 7) confirm that there is a significant difference in the mean scores of the two different samples, at the 5 % level. The South African sample generally performs at 7.5 points below the British sample. It is interesting to note the differences in Standard deviations in the two samples. There is more variation around the mean score in the British sample (SD=6.24), whereas the South African sample displays less variation around the mean per age group (SD=1.92). This SD has changed from 5.73 on the total sample to 1.92 when omitting the youngest and oldest age groups.

Table 7

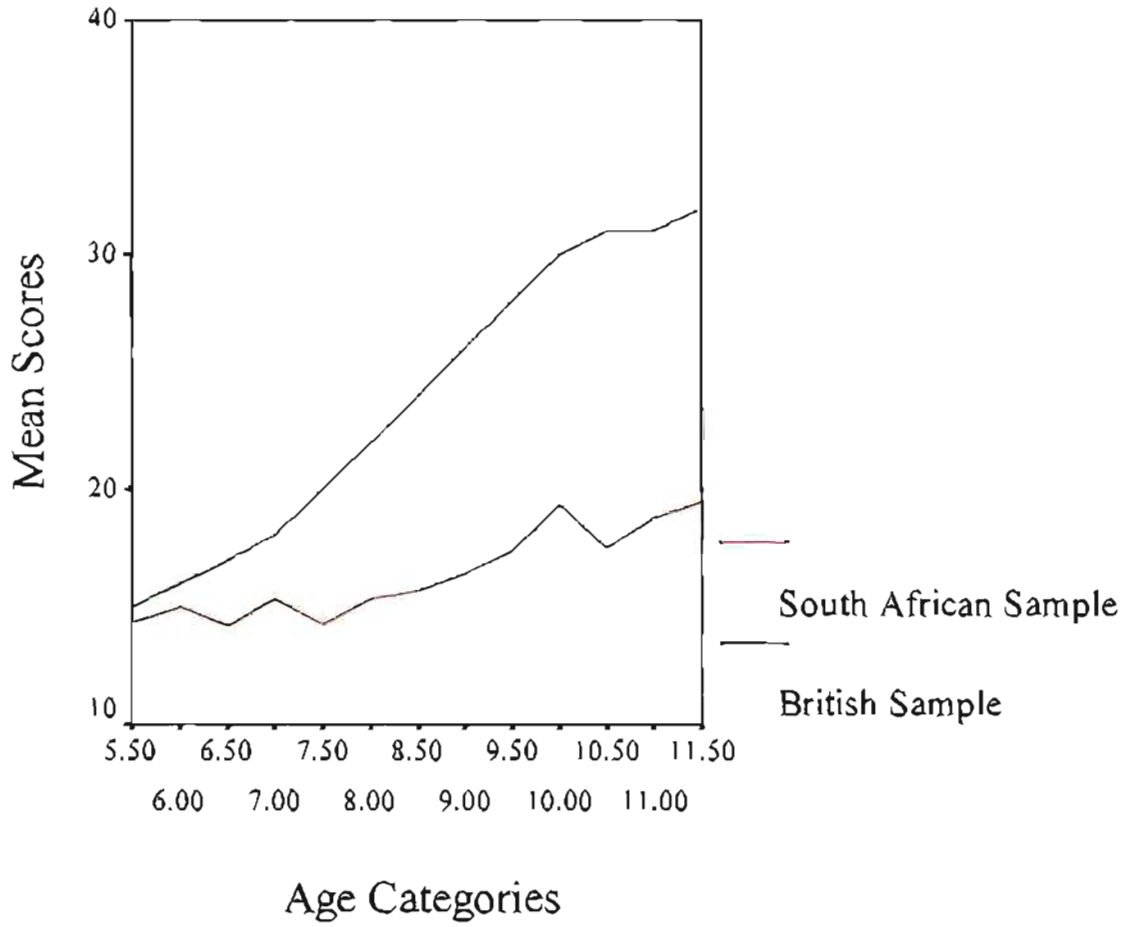
t-test results for the comparison of British and South African samples on individual scores

Population	*t*-value	Degrees of Freedom	Significance
British/South Africa	4.139	24	0.000

The graph in Figure 2 provides a good visual demonstration of the differences in performance across the two samples.

Figure 2

Plot of mean scores by age per sample



The above results confirm Hypothesis 1. There is a significant discrepancy, taking age into account, between the performance on the RCPM of this study's South African sample and the British standardisation sample. The results, furthermore, indicate the direction of this difference, with the South African sample performing well below that of the British (mean difference = 7.5 and SD difference = 4.3).

4.3 COMPARISON OF SCORES ACROSS THE INDEPENDENT VARIABLES

The participants' scores were analysed in relation to age, gender, grade and geographical location.

4.3.1 Hypothesis 2: Comparison of Scores across Ages

The mean scores for the sample by age are presented in Table 8. The age interval of 13 years is not shown in this table, since there was only one participant who fell within this category. The graph in Figure 3 shows small peaks and valleys. A steady increase across ages is not evident.

Figure 3

Plot of mean scores by age

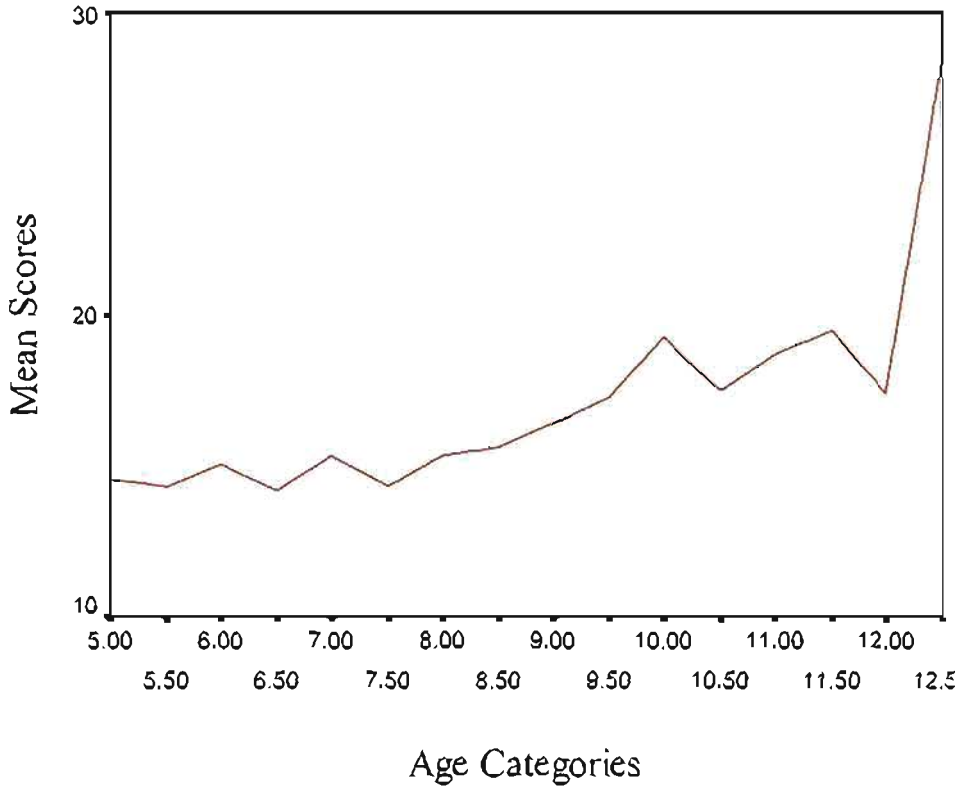


Table 8

Mean Scores, Standard Deviation, Range, Minimum and Maximum Scores by Age

Age	n	Mean	SD	Min	Max	Range
5	2	14.5	0.71	14	15	1
5.5	3	14.3	3.79	10	17	7
6	24	15	4.94	4	26	22
6.5	23	14.2	4.6	6	24	18
7	32	15.3	4.85	6	25	19
7.5	40	14.3	4.77	5	23	18
8	46	15.3	4.46	8	33	25
8.5	53	15.6	4.51	7	25	18
9	64	16.4	5.81	5	33	28
9.5	54	17.3	5.83	7	32	25
10	65	19.3	6.46	8	31	23
10.5	50	17.5	5.44	6	27	21
11	33	18.7	6.79	6	29	23
11.5	22	19.5	6.81	10	31	21
12	8	17.4	4.69	13	26	13
12.5	2	28.5	3.54	26	31	5

The results from the multiple regression analysis, presented in Table 13, indicate that age has a significant effect on RCPM performance, despite the observation that performance does not increase steadily across the age groups. Hypothesis 2 is thus confirmed.

4.3.2 Hypothesis 3: Comparison of Scores across Grades

The mean scores of the sample by grade are presented in table (Table 9) and graph (Figure 4) format.

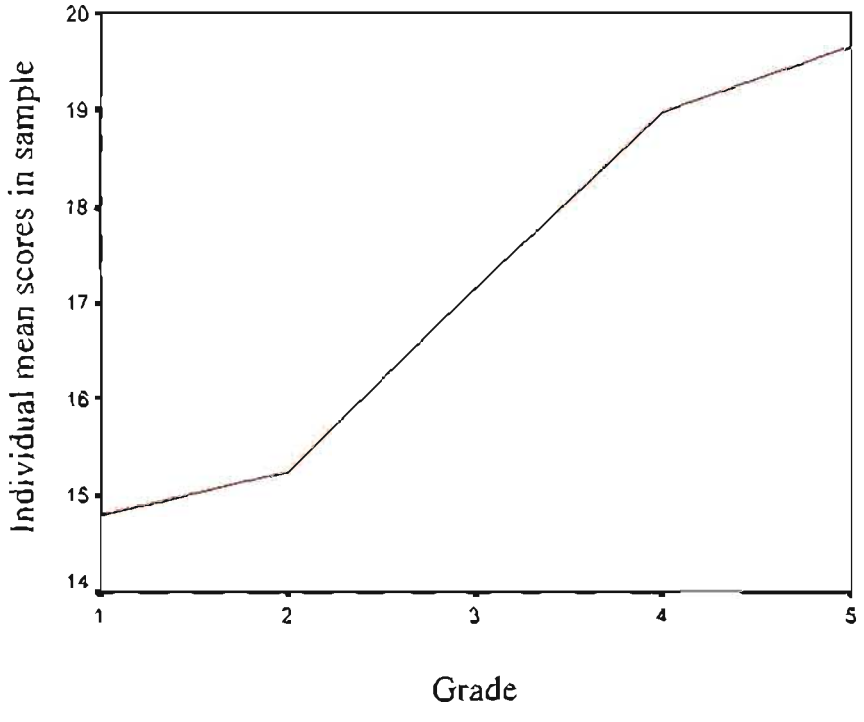
Table 9

Means, standard deviations, range and minimum and maximum scores by grade

Grade	n	Mean	SD	Min	Max	Range
1	133	14.8	4.58	4	26	22
2	114	15.2	4.98	7	33	26
3	131	17.2	5.39	5	32	27
4	101	19	6.63	6	33	27
5	43	19.6	6.4	8	31	23
Total	522	16.7	5.73	4	33	29

Figure 4

Plot of mean scores by grade



It is readily apparent from the graph, that there is an increase in the mean CPM scores across the grades. There appears to be a steep incline in performance from Grade 2 to Grade 4. Post-hoc comparisons of means using Levene's HSD for unequal sample sizes, revealed significant differences at the 5 per cent level, between all grades except 1 & 2 and 4 & 5. (See Table 2 in Appendix E for the results of the post-hoc comparisons).

The level of education, as represented by grade has a significant effect on performance of the RCPM. As grade increases, so do individual scores, indicating a positive correlation between educational level and performance. This confirms Hypothesis 3.

4.3.3 Hypothesis 4: Comparison of Scores across Gender

The mean score for the male participants in the sample was 17.3, with a standard deviation of 6.21. The female participants in comparison scored 1 point below the male mean, with a mean of 16.1 and a standard deviation of 5.14. The maximum score achieved by the male participants was 33 and the minimum was a score of 4. The female participants performed within a similar range, displaying a minimum score of 6 and a maximum score of 32. The results from a *t*-test performed on the mean scores of both genders, are presented in Table 10.

Table 10

t-Test results for the comparison of scores across Gender

Gender	*t*-value	Degrees of Freedom	Significance
Male/Female	-2.419	520	0.016

The results of the *t*-test confirm a difference between the mean scores of male and female participants in the sample and indicate that this difference is significant at the 5 per cent level. Males therefore tend to score significantly higher than females in this sample.

The above results do not confirm Hypothesis 4. Gender does seem to have a significant effect on performance, where males tend to outperform females.

Tables 11 and 12 show the means and standard deviations for males and females by grade and age respectively.

Table 11

Mean Scores and Standard Deviation for Males and Females by Grade

	Gender					
	Female			Male		
Grade	n	Mean	SD	n	Mean	SD
1	58	14.7	3.73	75	15.3	5.26
2	55	14.4	5.02	59	15.9	5.45
3	70	16.6	5.48	61	17.8	5.87
4	52	17.4	6.62	49	19.4	7.57
5	24	17.2	5.44	19	22.6	6.16
Total	259	16.1	5.14	263	17.3	6.21

Table 12

Mean Scores and Standard Deviation for Males and Females by Age

	Gender					
	Female			Male		
Age	n	Mean	SD	n	Mean	SD
5	0	0	0	2	10.5	2.12
5.5	2	13.5	4.95	1	16	
6	9	15.6	3.84	15	15.3	5.23
6.5	14	13.5	3.72	9	14.9	5.42
7	19	15.1	4.5	13	16.2	6.33
7.5	20	13.9	5.42	20	14.2	5.27
8	19	15.4	3.32	27	15.4	5.32
8.5	26	15.4	5.86	27	15.3	4.59
9	30	16.1	5.01	34	16.9	6.40
9.5	30	15.9	6.56	24	18.5	5.63
10	33	17.2	5.88	32	21.0	7.04
10.5	25	16.1	4.99	25	18.8	5.56
11	17	17.9	6.42	16	20.0	6.84
11.5	8	16.6	5.66	14	21.7	6.65
12	5	16.6	2.30	3	11.3	10.60
12.5	2	28.5	2.12	0	0	0
13	0	0	0	1	12.0	0
Total	259	16.1	5.14	263	17.3	6.21

4.3.4 Hypothesis 5: Comparison of Scores across Geographical location

The sample was divided into two different locations, namely rural and urban. The means and standard deviations of the individual RCPM scores for these two sub-samples are shown in Table 13.

Table 13

Means and standard deviations of scores by geographical location

Location	n	Mean	SD
Rural	285	16.14	5.741
Urban	237	17.35	5.658

The urban sample achieved a higher mean score of 17.35 in comparison to the rural sample's mean score of 16.14. A *t*-test was performed to investigate the significance of this difference. Table 14 shows that there is a significant difference at the 5 per cent level between the two group's mean scores, with the urban sample mean scores significantly higher than their rural counterpart's.

Table 14

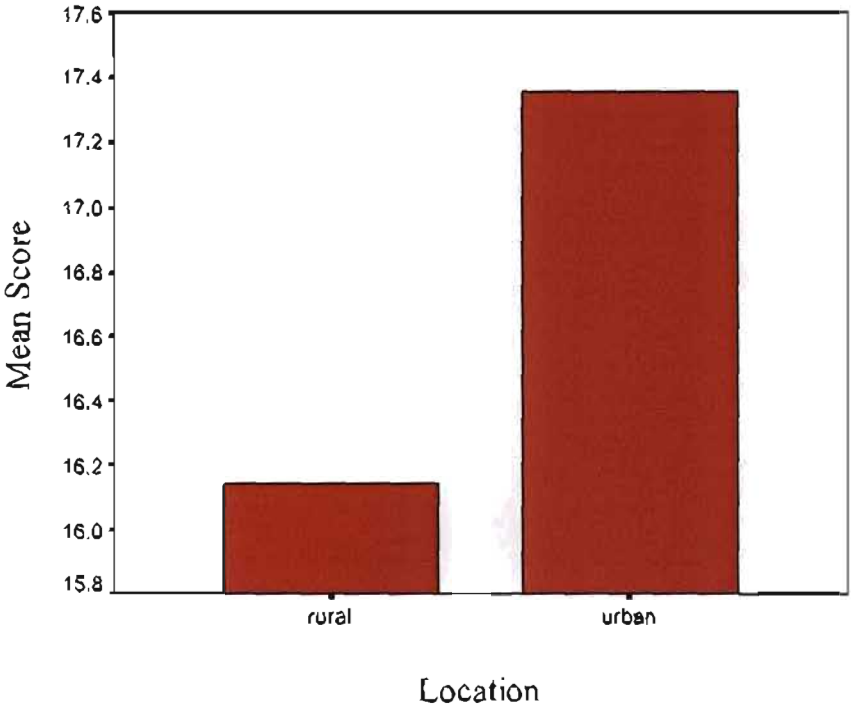
t-test Results for the comparison between Rural and urban Samples Mean Scores

Location	*t*-value	Degrees of Freedom	Significance (2-tailed)	Mean Difference
Rural/Urban	-2.397	520	.017	-1.20

The direction of these results confirms the expectation that the urban sample would outperform the rural sample. Hypothesis 5 is therefore confirmed; geographical location has a significant effect on performance of the RCPM. Figure 5 presents the difference in a bar graph.

Figure 5

Difference in means between rural and urban sample



Tables 15 and 16 present the means and standard deviations for rural and urban participants by grade and age respectively.

Table 15

Mean Scores and Standard Deviation for Rural and Urban by Grade

	Location					
	Rural			Urban		
Grade	n	Mean	SD	n	Mean	SD
1	85	14.5	4.59	**48**	16.0	4.63
2	54	15.1	4.76	**60**	15.3	5.74
3	53	16.3	5.81	**77**	17.7	5.55
4	50	17.4	7.02	**51**	19.3	7.17
5	43	19.7	6.31	**1**	**14.0**	**0**
Total	285	16.1	5.85	237	17.3	5.97

Table 16

Mean Scores and Standard Deviation for Rural and Urban by Age

	Location					
	Rural			Urban		
Age	n	Mean	SD	n	Mean	SD
5	1	12.0	0	1	9.0	0
5.5	1	17.0	0	2	13.0	4.24
6	8	12.3	4.1	16	17.0	4.2
6.5	13	13.5	5.13	10	14.8	3.33
7	15	14.8	4.7	17	16.2	5.78
7.5	18	14.4	5.23	22	13.7	5.43
8	23	14.7	3.3	23	16.1	5.51
8.5	19	13.5	4.79	34	16.4	5.19
9	33	15.7	5.25	31	17.4	6.17
9.5	24	16.2	5.69	30	17.7	6.67
10	40	18.1	6.07	25	20.7	7.43
10.5	31	16.0	5.58	19	19.8	4.23
11	30	19.2	6.54	3	16.0	8.00
11.5	20	19.4	6.75	2	25.0	2.83
12	7	14.6	7.04	1	15.0	0
12.5	2	28.5	2.12	0	0	0
13	0	0	0	1	12.0	0
Total	285	16.1	5.74	237	17.4	5.66

4.3.4 Multiple Regression Analysis

It is clear from the results of the t-tests run for all independent variables, that they all have a significant effect on individual test scores (dependent variable). A multiple regression analysis was performed to assess the predictive value of each independent variable. Weight, height and BMI scores were included as covariates. Table 137 presents the results of the multiple regression analysis.

Table 17

Multiple Regression Analysis: Coefficients

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
(Constant)	11.270	23.466		.480	.631
age of subject	.540	.265	.149	2.035	.042*
grade subject is in	.907	.294	**.202**	3.090	.002*
gender of sub	1.398	.474	.122	2.948	.003*
individual weight of subject	.043	.367	.048	.118	.906
individual heights	-1.216	17.785	-.019	-.068	.945
bmi scores of sample	-.168	.678	-.063	-.247	.805
situated in rural/urban	1.697	.492	.148	3.452	.001*

* denotes significance at the 5% level

As expected, age, grade, gender and geographical location are all significant predictors of performance on the RCPM. The most important predictor of performance on the RCPM appears to be grade, as indicated by the highest Beta (ß) value of .202. It is interesting to note that the individual weight, height and total BMI scores0 are not related to the individual scores on the sample.

4.4 RELIABILITY

The 36 items of the RCPM produced a very respectable Cronbach's Alpha of .85, which indicates good internal consistency for the scale on this sample. The items appear to correlate very well to the total scale. Table 18 shows item-total correlations and the resulting alpha if items are deleted. There are a few items on the test, which seem to have lower correlation coefficients than the rest of the scale items. These item-total correlations are highlighted. Removing most of these items however, does not improve overall reliability (Removing B12 will give an increase of 0.02). Item Ab12 however, has a very low negative correlation, and Cronbach's alpha would be increased by 0.05 if this item is deleted.

Table 18

Cronbach's alpha for thirty-six items of the RCPM

Item	Item Total Correlation	Alpha if Item Deleted
A1	.217	.849
A2	.216	.849
A3	.254	.848
A4	.303	.847
A5	.352	.846
A6	.385	.845
A7	.348	.846
A8	.184	**.851**
A9	.298	.848
A10	.393	.845
A11	**.158**	**.851**
A12	**.175**	.850
Ab1	.355	.846
Ab2	.408	.844
Ab3	.342	.846
Ab4	**.541**	.840
Ab5	.473	.842
Ab6	.508	.841
Ab7	.562	.840
AB8	.571	.840
Ab9	.360	.846
Ab10	.409	.844
Ab11	.267	.848
Ab12	**-.073**	.855

Item	Item Total Correlation	Alpha if Item Deleted
B1	.395	.845
B2	.449	.843
B3	.558	.840
B4	.514	.841
B5	.510	.841
B6	.409	.844
B7	.250	.848
B8	.202	.849
B9	**.188**	.849
B10	**.181**	.850
B11	**.180**	.849
B12	**.111**	.852
Total Cronbach's Alpha		**.85**

4.5 VALIDITY

4.5.1 Item Analysis

The item analysis also reveals useful results for measures of validity. Low item-total correlations as highlighted above, indicate items which appear to be poor at discriminating between subjects' performance on the test. These items are A11, A12, Ab12, B9, B10, B11, B12. The item Ab12, in particular, is very poor with a negative item-total correlation of -.073.

Table 19 presents the responses per item for the total sample. Firstly, it is apparent that not all children answered A1 and A2 correctly. (These two items are used as practice examples, implying that all children should answer them correctly) The percentage of correct responses is high at the beginning of each set and then gradually decreases. For Set Ab and Set B, the drop in correct responses seems to be larger after the first two items, followed by a gradual decline, when compared to Set A. It is however, clear that a few items appear to be out of sequence.

These items have been highlighted and are identified as A7, Ab10 and Ab11. Only 29 percent of the sample chose the correct response to A7. This low percentage is marked by its position in the sequence of items, where the preceding item (A6) was correctly responded to be 79 percent of the sample and the following item (A8) was correctly answered by 58 percent of the sample. In contrast, Ab10, Ab11 and B12 have more correct responses than their preceding and proceeding items.

Table 19

Proportion of correct and incorrect responses per item

Item	No. of Correct Responses	% Correct Responses	No. of incorrect Responses	% Incorrect Responses
A1	513	98.3	9	1.7
A2	**504**	96.6	18	3.4
A3	**491**	94.1	31	5.9
A4	**477**	91.4	45	8.6
A5	**415**	79.5	107	20.5
A6	**412**	78.9	110	21.1
A7	**153**	**29.3**	**369**	**70.7**
A8	**304**	58.2	218	41.8
A9	**248**	47.5	274	52.5
A10	**221**	42.3	301	57.7
A11	**118**	22.6	404	77.4
A12	57	**10.9**	465	89.1
Ab1	**449**	86	73	14
Ab2	**414**	79.3	108	20.7
Ab3	**417**	49.9	105	20.1
Ab4	**222**	42.5	300	57.5
Ab5	**200**	38.3	322	61.7
Ab6	**165**	31.6	357	68.8
Ab7	**199**	**38.1**	323	61.9
Ab8	**133**	25.5	389	74.5
Ab9	**110**	21.1	412	78.9
Ab10	**180**	**34.5**	**342**	**65.5**
Ab11	**141**	**27**	**381**	**73**
Ab12	**74**	14.2	448	85.8

Item	No. of Correct Responses	% Correct Responses	No. of incorrect Responses	% Incorrect Responses
B1	424	81.2	98	18.8
B2	343	65.7	179	34.3
B3	256	49	266	51
B4	236	45.2	286	54.8
B5	**169**	32.4	353	67.6
B6	**146**	28	376	72
B7	**106**	20.3	416	79.7
B8	**59**	11.3	463	88.7
B9	**61**	11.7	461	88.3
B10	**93**	17.8	429	82.2
B11	**42**	8	480	92
B12	**122**	**23.4**	**400**	**76.6**

4.5.2 Factor Analysis

Factor analysis was performed on the data collected from the total sample. Principal component analysis was the method chosen, with two different types of rotation used. Both a varimax and oblique rotation were used to transform the data. The eigenvalues and scree plot (Figure 6) suggest that there is one dominant underlying factor, which most items tend to load on. Thereafter the curve drops noticeably. Table 20 outlines the eigenvalues of the first three components and the percentage of variance, which they account for. (See Table 3 in Appendix E for the full detail on all components).

Figure 6

Scree Plot of principal component analysis. Thirty-six components are identified.

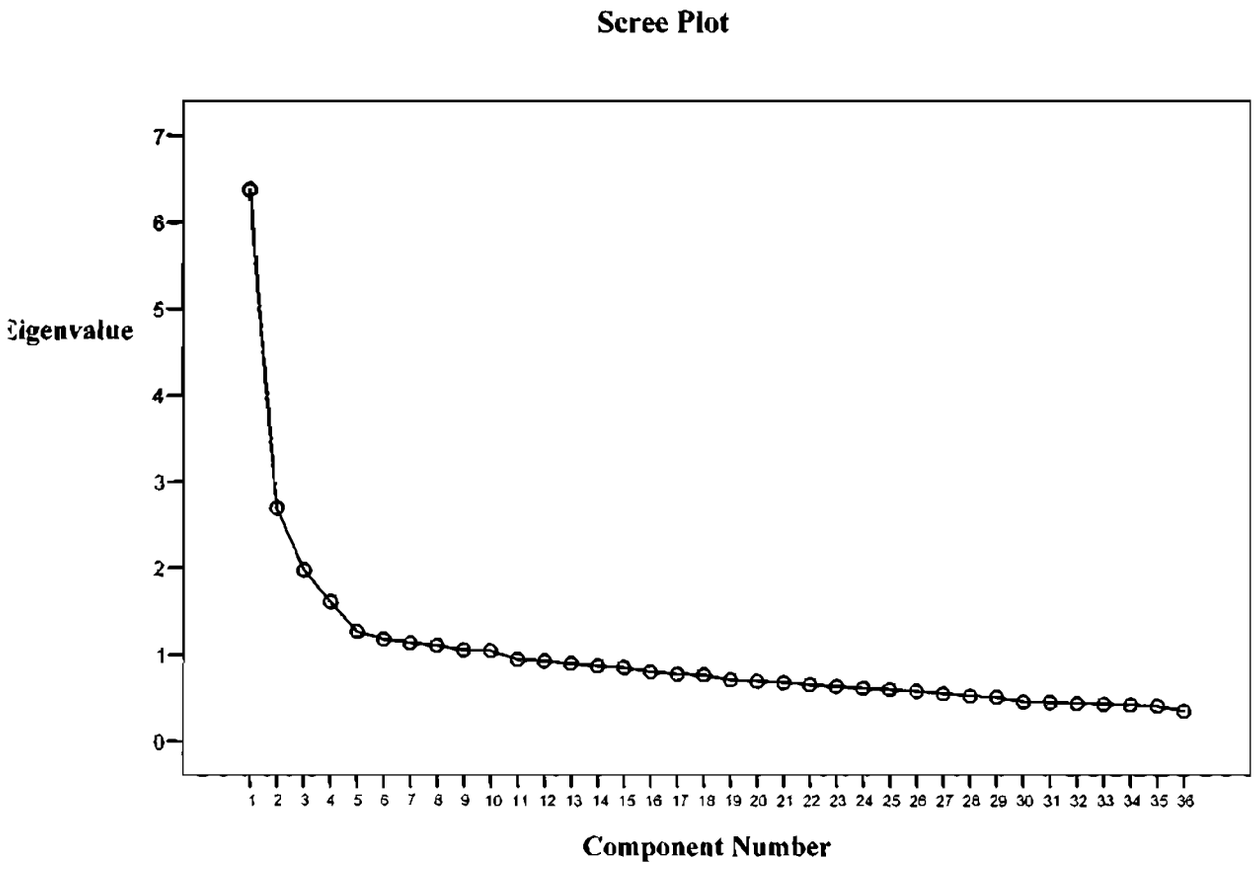


Table 20

Extracted factors: eigenvalues and percentage of variance

Factor	Eigenvalue	% Variance	Cumulative %
1	6.378	17.716	17.716
2	2.702	7.506	25.223
3	1.977	5.493	30.715

Factor 1 accounts for 17 percent of the total variance, which confirms its dominance in this factor structure. The first three factors together explain 31 percent of the total variance. Since the results of the two different types of rotation, were very similar in their identification of the types of items which load on the three factors, Table 21 presents only those results of the Principal component analysis with varimax rotation and Kaisser normalisation. (The results from this rotation have been chosen to facilitate comparison with other factor analytic studies on the RCPM.) Appendix E can be consulted for a similar table (Table 4) describing the results from the oblique rotation.

Table 21

Rotated Component Matrix for RCPM items and factor loadings

Item	Component		
	1	2	3
A1	-.022	**.474**	.118
A2	-.046	**.541**	.104
A3	.010	**.570**	.011
A4	.121	**.510**	-.052
A5	.178	**.495**	.068
A6	.199	**.543**	.008
A7	**.448**	.051	.057
A8	.175	.150	-.073
A9	.294	.183	.075
A10	.397	.194	.047
A11	.203	.007	-.007
A12	.356	-.188	-.039
Ab1	.111	**.650**	-.015
Ab2	.230	**.555**	.026
Ab3	.119	**.614**	-.046
Ab4	**.626**	.181	.022
Ab5	**.619**	.050	-.004
Ab6	**.638**	.071	.113
Ab7	**.670**	.105	.112
Ab8	**.706**	.088	.059
Ab9	**.479**	-.032	.137
Ab10	**.475**	.105	.057
Ab11	.329	.093	-.032
Ab12	.020	-.220	-.011
B1	.262	**.547**	-1.27
B2	.362	**.505**	-.140
B3	**.574**	.295	**.033**
B4	**.587**	**.206**	.019
B5	**.615**	**.077**	.124
B6	**.455**	.120	.153
B7	.274	-.015	.327
B8	.134	-.030	**.704**
B9	.095	.033	**.666**
B10	.039	.092	**.664**
B11	.073	.026	**.736**
B12	.165	.050	-.138

Note: loadings ≥ .40 considered relevant for interpretation of factors.

Factor 1 is described by items A_7 (simple continuous pattern completion), Ab_4, Ab_5, Ab_6, Ab_7, Ab_8, Ab_9, Ab_{10}, B_3, B_4, B_5 (discrete pattern completion of related figures as a whole), B_6 (concrete reasoning by analogy). Items with significant loadings on Factor 2 are A_1, A_2, A_3, A_4, A_5, A_6 (simple continuous pattern completion), Ab_1, Ab_2, Ab_3 (discrete pattern completion involving identity), B_1, B_2 (discrete pattern completion). Factor 3 is marked by B_8 B_9 (concrete reasoning by analogy) and B_{10}, B_{11} (abstract reasoning by analogy). Inspecting the number of items and those which had the highest loadings on the factors, the factors could be interpreted as follows: Factor 1 could be *Discrete pattern completion of related figures as a whole*; Factor 2, *Simple continuous and discrete pattern completion* and Factor 3 could be defined as *Concrete and abstract reasoning by analogy*.

(All SPSS output is available on request.)

The implications of the above results will be explored further in the next chapter.

CHAPTER FIVE

DISCUSSION

In this chapter the results presented in chapter four will be discussed in terms of the aims and hypotheses of this study and the literature reviewed in chapter two. The limitations of the study will be examined and the implications of the present findings for future research will be explored.

5.1 NORMATIVE EVALUATION

The results in section 4.2 confirm Hypothesis 1, in that there was a significant discrepancy between the performance of this study's sample and the British standardisation sample. This local sample's mean performance was substantially lower (by 7.5 points) than that of the British standardised sample's performance. (This could be equated with two and a half years below children in the British sample).This is consistent with other findings from Africa, both on the RCPM and RSPM, where performance is substantially lower than European norms (Costenbader & Ngari, 2001; Jinabhai et al., 2004; Bass, 2000; Owen, 1992; Notcutt, 1949; Zindi, 1994; Vass, 1992).Both early (Notcutt, 1949) and more recent studies (Jinabhai et al., 2004) have noted a trend in Zulu children to perform at three to five years below Euro-American norms on the RCPM. Lower performance has also been noted in a Xhosa sample of children (Bass, 2000) and across 46 different tribal groups in Kenya (Costenbader & Ngari, 2001).

5.1.1 Socioeconomic Status (SES)

Other reports of lower norms have been directly linked to socioeconomic status (SES) in the United States of America (USA). Findings suggest that lower SES results in poorer performance and that when different groups are matched for SES, their performances are more comparable (Emerling, 1990; White & Panunto, 1978). Given the observations made during the study, this study's sample of Zulu children, both rural and urban sub-samples, could be considered

representative of a much lower SES group than that used in the British standardization sample. Different levels of socioeconomic status could therefore be partly responsible for the discrepancy between these two samples' performances.

5.1.2 Nutrition

It is well known that SES is associated with nutritional status, where poorer people have less access to food sources and their food is of poor quality (Whaley et al., 2003). Poverty and malnutrition are common in developing countries. One of the interventions aimed at addressing this deficit in South Africa, are feeding schemes. These are set up in schools in poorer areas to provide children with one balanced meal in their day. Both schools in the sample for this study benefit from feeding schemes. Recent research in Africa, however suggests that supplementation in itself is not necessarily sufficient, but animal food source protein is necessary for optimal cognitive performance (Whaley et al., 2003). The meals provided by the feeding schemes in this sample school, consisted mainly of samp and beans. There seemed to be either an absence or very limited amount of animal protein in the diet (N. Mzila, personal communication, November, 2004). Although the results of the relationship between BMI and performance for this sample were not significant, the amount of information which this measure of nutritional status is able to provide is considered limited. These limitations are explored under 5.4. It is suggested that despite the results, poor quality nutrition may play a role in the lowered nonverbal performance measured by the RCPM of this study's sample. Furthermore, improved nutrition in developed countries, like Britain, has been correlated with an increase per decade, in nonverbal performance in children (Lynn, 1989; Raven et al., 1998). Not only may the Zulu children have poorer nutrition than the British sample, but the British sample seems to have the accumulated benefit of improved nutrition over the past few decades. This would seem to widen the gap between the two samples' performances.

5.1.3 Culture

One of the most plausible accounts of the differences observed between these two samples is the role of culture. Although culture is a complex concept, as explored in the literature review, for

the purposes of this discussion, the notion of cultural loading is most relevant. Van de Rijt (1990, as cited in Helms-Lorenz & Van de Vijver, 1995) showed how the difference in performance on cognitive tasks, between two different cultural groups, increased as the cultural loadings on the tests increased. Cultural loadings, which may be identified on the RCPM in this study, include the following aspects of assessment: tester, test-takers, interaction between tester and test-takers, response protocols and test items. While subject related factors and test item aspects are suggested as contributing the most to the differences observed between these two samples, subject related factors will be explored here, and the impact of test items will be discussed later in connection with the results on validity.

Although Raven et al. (1998) claim that the RCPM is not dependent on verbal ability, nor is it necessary to verbalise instructions for the test, this has been contradicted by other research. Researchers using the RCPM in Africa have frequently commented on the need for simplified and repeated instructions for African pupils (Wober, 1969). This was observed during the administration of the test in the current study. This included understanding the response protocol format. The present study sample found the RCPM answer format unfamiliar and needed additional instruction in how to record responses. It appears then that African children's verbal ability and attention is critical to their understanding of how to complete the RCPM. This suggests that nonverbal measures may still require verbal ability for completion, particularly for understanding the test instructions. Difficulties with an unfamiliar test format are considered here as related to the broader difficulties encountered with an unfamiliar experience of testing. The increased arousal that was observed in most of the children due to the novelty of the situation (external researcher of different culture, using new books) may have been a distracter in performance. Concentration and motivation to perform on tests is related to the subject's degree of test-wiseness. Test-wiseness is considered less sophisticated in non-western cultures (Nell, 1999).

5.1.4 Perceptual Differences

Not only does familiarity with the test format appear to have an effect on performance, but so too may familiarity with the test content play a role in this sample's performance. The RCPM

consists of problems designed in an abstract format, using pictures of shapes and symbols (Raven et al., 1998). Items on the RCPM seem to rely on visuospatial competency (Lezak, 1995). Research has shown that with African subjects, familiarity with pictorial material is critical to their perception of that material (Berry, 1971; Serpell, 1979). There are further suggestions that direct experience with pictures is necessary (Hudson, 1962; Deregowski, 1980; Miller, 1973; Serpell, 1979). Poorer performance on tasks of abstract representation should not be interpreted as an inability to use abstract concepts, but as a difficulty with an unfamiliar medium of representation (Serpell, 1979). It is suggested that many of this sample have had little access to pictorial material, in the form of books, magazines, or drawings. Both schools had no libraries and textbooks were limited. Children often see and use books for the first time in school (N. Mzila, personal communication, August, 2005). Difficulties perceiving the RCPM problems could have thus affected this sample's performance.

The improved performance seen on second testing of African subjects with the RCPM seems to confirm the findings that test-wiseness and familiarity with the medium of test content are influential factors in performance (Wober, 1969). It is likely that the subjects have had an opportunity to learn test-taking skills and use a different medium of representation. On second exposure to testing, they therefore know what is required and are able to work accordingly. Wober (1969) suggests that the coaching gain provides a more valid reflection of the African subject's ability. A dynamic assessment procedure may therefore be more appropriate, where subjects have an opportunity to become familiar with both the test content and testing experience.

5.1.5 Education

In the results, it is interesting to note that although there is an overall increase in performance across ages in the Zulu sample, the gap between their performance and the English sample's performance seems to widen with age. This suggests that education is an important factor in the observed differences. It further highlights the differences that may exist in these two culture's educational systems. Ardila and colleagues (1989) have proposed that education is a stronger moderating variable than culture in cognitive performance, based on research showing that

cognitive abilities are highly trained skills. The educational syllabus used in the Zulu sample's schools may not therefore, be providing the specific experience or training necessary for promoting the development of abstract reasoning skills, as measured by this instrument.

This contradicts the belief of Raven et al. (1998) that all children will develop the ability of abstract reasoning by the age of 11 years. Raven et al. (1998) based their conception and design of the RCPM on Piaget's theory of cognitive development. Piaget believed that the development through different stages was age-dependent and universal across all cultures. Recent research however, suggests that specific experiences are necessary to learn new concepts and to proceed through the successive stages (Berk, 2000). Culture and training both therefore seem to be implicated in cognitive development. Vygotsky recognized the role that both culture and education play in his theory of cognitive development. He believed that the values of a child's society, as transmitted through culture and education, interacted with genetic predispositions to determine cognitive development (Thomas, 1985). The different performances across the two cultures seem to be better explained by Vygotsky's theory of cognitive development, than by Piaget's. If results were interpreted in light of Piaget's theory, it would have to be concluded that the Zulu children are deficient in abstract reasoning ability and differences in performance are real differences in nonverbal intellectual functioning.

The Universalist position adopted in this study, rejects such an explanation. It is suggested here, and supported by evidence, that the difference in performance across these two different cultural groups is the result of an interaction between many factors. Culture, perceptual skills, SES, nutrition and education all appear to influence test performance. The significant difference in the performances of these two samples highlights the need for local normative data. These findings show that by using the British norms for a Zulu child, intellectual performance is likely to be interpreted as deficient. The need for local norms has been emphasized by both researchers within South Africa (Anderson, 2001; Nell, 1999; Shanahan, Anderson, & Mkhize, 2001; Shuttleworth-Jordan, 1996; Viljoen et al., 1994) and those involved with cross-cultural psychological research (Ardila, 1995; Geisinger, 1994; Helms-Lorenz & Van de Vijver, 1995; Sigmon, 1983). It must also be noted that the distribution of scores for this sample is not normal. Use of the median may be considered more meaningful, as the mean is a measure of central

tendency and the standard deviation explains the variance around that central tendency (Hammond, 1995; Terre Blanche, 2005). The raw scores from this sample may therefore be problematic to use in compiling population norms.

5.2 IMPACT OF INDEPENDENT VARIABLES

5.2.1 Age and Grade

The effects of age and grade on performance will be discussed together. Hypothesis 2 is confirmed by the findings that age has a significant effect on performance. An increase in scores is correlated with an increase in age. In order to understand the impact of age on performance, it needs to be considered in conjunction with the effect of educational level on scores. The valleys observed in the graph for some age categories may be explained by the lower grades which some of these participants fell into. The results of the multiple regression analysis indicate that the grade the child is in is the most significant predictor of performance. This adds to the results discussed above, where an increasing difference was found in performance across the two cultural samples, with age. The argument has been made that if Piaget's theory was valid, this difference would not be seen. It is suggested that educational level is thus more significant. Indeed, results show an increase across grades, thereby confirming the third hypothesis. This increase does not follow a predictable pattern in this study. There is a sharp increase in performance from Grade 2 to Grade 4, but not such marked improvements observed between Grades 1 and 2, 4 and 5. The reasons for this are not clear at this stage and require further exploration. A possible explanation has to do with the school curriculum and a change in the nature of the tasks expected of learners between Grades 2 and 3. The findings regarding the impact of grade level, support Ardila et al.'s (1989) hypothesis that nonverbal cognitive skills are highly trained and therefore dependent on education. The specific experience hypothesis in perceptual skills as proposed by Serpell (1979) and Berry's (1971) ecological demand model, provides additional explanation for the impact that education would have on visuospatial performance.

The possible impact of acculturation may also be related to the increase of performance observed across grades. Acculturation has been shown to be positively correlated with cognitive performance. Formal schooling is considered a powerful acculturating factor, where the more educated the individual is within that school system, the more acculturated to the dominant culture he becomes (Berry et al., 1992).

Considering that there was a large range of age groups observed in each grade, and that grade appears to be the most significant predictor of performance on the RCPM, local normative data should possibly be organised according to educational level for Zulu speakers, rather than age categories.

Finally, given these findings, where educational level is indicated as the strongest predictor of performance on the RCPM, it can be concluded that for this sample, the RCPM cannot be considered an example of a "pure" aptitude test (Helms-Lorenz & Van de Vijver, 1995). It seems that the test measures skills which are acquired, where aptitude tests are supposed to measure skills which are less dependent on learning.

5.2.2 Gender

It was hypothesised that gender would not have a significant effect on performance. The results however, showed that males tend to perform better than females on the RCPM. Hypothesis 4 was thus not confirmed. These findings have been reported in other cross-cultural studies (Klippel, 1975; Sigmon, 1985) and more importantly are consistent with the reports in African studies (Bass, 2000; Costenbader & Ngari, 2001). The differences observed between the genders have often been explained in terms of culture, where societies are organised according to strong traditional gender roles (Berry, 1971; Costenbader & Ngari, 2001). The Zulu culture is one such culture where these divisions in roles are still apparent. The degree of acculturation is likely to affect how closely these gender roles are adhered to. Typically westernised societies display less differentiation in gender roles.

Research exploring the relationship between gender and cognition has repeatedly found that men outperform women on a variety of spatial tasks. Theories have focused on human evolution and the natural selection of abilities associated with an evolutionary history of gender specific tasks (Kimura, 2000). Gender differences in performance on the RCPM can therefore be explained by biological determinants and sociocultural patterns which promote the acquisition of different cognitive skills.

5.2.3 Location

The results confirmed the expectation outlined in Hypothesis 5, that the urban sample would perform significantly better than the rural sample. Rural areas have shown poorer performance in other studies (Raven et al., 1998; Sigmon, 1983). Within this sample, acculturation is likely to play a large role in the differences observed between these two groups. It has already been shown that the impact of acculturation can be defined in terms of different geographical locations (Shuttleworth-Jordan, 1996) and that this is related to access to resources (Nell, 1999). In terms of school quality and access to resources, there were not many differences observed between the two schools in the urban and rural areas. Teacher-student ratios appeared similar and information on the level of teacher training was not collected. Both schools had no library and although the urban school appeared to make more use of textbooks in teaching, this was merely an observation during the period of testing. Focus is therefore turned to other factors associated with the different locations, like test-wiseness, parental education and nutrition.

Urbanisation and access to the media are strong acculturative indicators (Berry et al., 1992), which have been linked to a greater degree of test-wiseness (Nell, 1999) and increased cognitive performance (Berry et al., 1992). It is further suggested that given the larger range in parental occupational functioning in the urban area, the level of parental education or skill may be related to the higher level of cognitive performance observed in the urban child. This may also be linked to the style of parenting, where the more educated parent may adopt a more authoritative style, which encourages field-independence in perceptual abilities (Dawson, 1967, in Miller, 1973; Witkin, 1962, in Berry, 1971). The role that parental education and style may play in a child's cognitive performance has been implied in Vygotsky's theory of cognitive development. He

believed that the skills valued by societies would be transmitted through socialising agents like parents, to the child (Luria, 1976, in Thomas, 1985).

Finally, the effects of nutrition on cognitive performance, which have already been highlighted, may be an explanation for differences between the two samples' performance. It is suggested that the urban children may have had better pre- and postnatal nutrition, than their rural counterparts. Pre- and postnatal care is likely to have been also improved due to better access to health care facilities.

The possibility of different norms for urban and rural children thus needs to be considered.

5.3 PSYCHOMETRIC PROPERTIES OF THE RCPM

The results of the reliability and validity analysis will be discussed and used to explore the question of test appropriateness for the study's sample.

5.3.1 Reliability

Reliability analysis revealed very satisfactory results. The high Cronnbach's alpha of .85 obtained, is consistent with the range of reliability coefficients reported in international and more specifically cross-cultural studies. With specific reference to African studies, the results are very similar to those reported by Costenbader and Ngari (2000) of .87, in Kenya; Kendall (1976) and Bass (2000) in South Africa, who found reliabilities of .87 and .88 respectively. The results suggest that this study sample interpreted and answered the items on the RCPM in a manner which is consistent with the intention of the design of the test. Factors which are likely to have contributed to a good reliability are the number of items on the test, the increase in difficulty of the items and the heterogeneous nature of the sample used in this study (Finchilescu, 2005; Wolfaardt, 2001). Although there was concern over the effect of changes in administration procedures from the original Raven's, these influences on the reliability of the measure, may actually be minimal.

5.3.2 Validity

The extraction of three principal factors by other authors (Carlson & Jensen, 1980; Schmidtke & Schaller, 1980; Wiedl & Carlson, 1976) is consistent with the current study's results. The factors identified are also considered similar. The factors identified in this study's analysis have been defined as Factor 1: Discrete pattern completion of related figure as a whole; Factor 2: Simple continuous and discrete pattern completion; and Factor 3: Concrete and abstract reasoning by analogy. Factor 2 and 3 were similarly defined and labelled as Factor 2 and 1 respectively by Wiedl and Carlson (1976). The definition of the factor structure identified in this study's analysis also confirms that the test is less differentiated in terms of the skills each factor draws on, than is implied in the test design. Nevertheless, the three factors appear to represent three different conceptual dimensions of the test. These results combined with good inter-item correlations from the reliability results, may seem to suggest that the test is tapping into a common 'g' factor. The findings seem to support Geisinger's (1994) review where similar factors have been reported across cultures, but the loadings of items have been different. Whether this factor structure is measuring the same construct in a different culture is debatable.

Assessing how well the test measures these identified dimensions, requires an investigation into item bias for this sample. The difficulty in analysing DIF has already been noted in chapter three. The proportion of correct responses per item has been used to gain a superficial view of DIF.

Set A contains the easiest items in the test and the results show that the greatest proportion of correct responses were indeed on the items in this first set. Items A1 and A2 should have 100 percent correct response since they represent the practice examples for the test. This is not shown, only 98 and 97 percent, respectively, of the sample answered these items correctly. It is suggested here that this response is more a reflection of difficulty with completing the answer protocol correctly than with not understanding the matrix problem. As expected, the first few items in each set are the easiest and these all have a high percentage of correct responses. These results show that this sample did interpret these items as Raven intended. There are however, a

few items which appear to be out of sequence: A7, Ab10, Ab11 and B12. Item A7 appears to be more difficult for this sample. Interestingly this finding was also reported in the Kenyan standardisation sample (Costenbader & Ngari, 2001). The item is supposed to measure simple pattern completion involving the ability to form a gestalt (Raven et al., 1998). Ab10, Ab11 and B12 all appear to be less difficult than preceding items in the sequence, for this sample. Both Ab10 and Ab11 measure discrete pattern completion of related figures as a whole. Ab10 requires the perception of open symmetry and orientation, while Ab11 uses closed asymmetry and oblique orientation. B12 is a measure of abstract reasoning involving a double subtraction process (Raven et al., 1998).

A superficial assessment of item validity would expect a uniform increase in the order of item difficulty. Considering the findings that a few items are not sequenced appropriately for this sample, there is an indication that rank ordering may need to be reviewed and adjusted for this sample.

Item-total correlations have revealed those items which appear to have a poor discriminatory power. They have been identified as items A11, A12, Ab12, B9, B10, B11 and B12. With a negative low correlation of -.073, Item Ab12 has the worst discriminatory power compared to all other test items. It must be noted that these items represent some of the most difficult items on the test according to its design (Raven et al., 1998) and the items which represent the lowest proportion of correct responses for the present sample. Overall, the positive skew of the distribution of the scores for this sample, suggests that more students performed poorly on the test than those who performed well. Combining all the results from item analysis with those obtained from the frequency distribution of scores, it is suggested that for most of these students, this test was too difficult. Item selection for this sample may therefore also require review and adjustment.

Using Owen's (1998) criteria to evaluate the absence of construct bias for this particular sample, results indicate the following:

1. there are similar test reliabilities between this sample and other cultural groups
2. there is evidence of similar factor structures

3. rank orders of item difficulty behave differently for this sample
4. as do some of the items' discriminatory power.

The results suggest that the last two criteria cannot be met. Based on these results, it may be proposed that the RCPM is not a culture-fair instrument. The nonverbal nature of the test does not therefore imply reduced cultural loadings. This suggestion supports those cautions issued by Anastasi (1990), Crawford-Nutt (1976) and Wober (1969). There is the added difficulty in investigating whether the construct of nonverbal intelligence is actually being measured in this sample. This would depend on how intelligence is actually defined for this particular culture. It is valid that the RCPM measures specific skills, but whether these skills represent intelligent behaviour for this sample is questioned. Given these preclusions, it is not possible to make a definitive statement regarding construct validity of the RCPM for this Zulu sample.

5.4 LIMITATIONS OF THIS STUDY

Although the sample drawn for this study was relatively large, it was a convenience sample and should not be considered a true representation of the entire Zulu population. The generalisability of the study's findings is thus limited. The South African sample was matched for age in the comparison with the British sample, but was not matched in size for each age category, nor was it matched on other sample characteristics like socioeconomic status. Furthermore in selection of the two geographical sub samples, there was no matching for sample size in either age group or grade.

The design of this study was limited by the small number of context variables chosen for investigation. The use of geographical location as an indication of the degree of acculturation is supported by various authors in the area (Berry et al., 1992; Shuttleworth-Jordan, 1996). The dynamic nature of this concept however, makes it very difficult to operationalise and determine where individuals lie on its continuum. By choosing to operationalise it in terms of rural and urban locations may still be too general for isolating the actual factors which have the largest impact. This has limited the discussion of these findings to speculation around the role of certain acculturative factors. The operationalisation of the other context variable, nutritional status,

must also be considered limited in the amount of information it lends this study. The body mass index is recognised as a good indicator of nutritional status and is widely used in district hospital settings (Centres for disease control and prevention, 2005; J. Galliers, personal communication, September 12, 2005). It may however obscure a number of other factors which would be considered indicators of poor nutrition. For example, a child may have an age appropriate height and weight, but still have an inadequate or poor quality diet. Excluding important and useful context variables, like measures of SES and level of parental education, must be considered a serious short-coming in this particular study.

This study was also open to various confounding variables, which were not controlled for. The impact of the following factors must be recognised as particularly detrimental to cognitive performance: head injury, HIV infection, learning disorders, mental retardation, trauma and other emotional disturbances.

Schaie (1965, as cited in Breakwell, 1995) highlight the importance of the socio-environmental context at the time of data collection. This is especially relevant when age-related development is the focus of study. Thus the limited use of variables which measure this context, is a weakness in this study's design.

Several factors during administration procedures may have interfered with the results obtained in this study. The examiner was of a different race, culture and language to the test-takers. This has been shown to contribute to cultural loadings on tests (Van de Vijver & Poortinga, 1992, in Helms-Lorenz & Van de Vijver, 1995). It is hoped that these effects may have been reduced with the help of familiar persons (teachers) in the testing procedure. While the use of teachers as aids, had advantages, the disadvantages were the difficulties in controlling what the teachers communicated to the children. Although they were fully instructed before testing, the examiner had little proficiency in the Zulu language and it was therefore unclear whether teachers followed instructions appropriately.

In administration of the test to his standardisation sample, Raven used mostly group testing and conducted individual testing where needed, (Raven et al., 1998). Although the procedure for

administration for the Zulu sample was similar in this regard, it is difficult to comment on the effect these two methods of testing may have had on individual performances. The examiner did not make a record of those who received individual administration and so it was impossible to investigate these effects. This is an unfortunate limitation. It was difficult to assess signs of stress or poor motivation in students during group administration of the test. There was also great excitement observed over the testing procedure, which may have obscured these signs. Group administration of the test therefore has its limitations.

Finally the time of day for testing was not consistent across all groups tested. Children are expected to be more alert in the morning and those tested at this time may have performed better, causing an unfair disadvantage for those tested towards the end of the school day. Time of testing may have therefore had an uneven effect on test performance.

5.5 IMPLICATIONS OF FINDINGS

When choosing to use a test designed outside the context in which its use is intended and for a population which is different from the original test population, the test must be evaluated for its suitability. Users of psychological tests have an obligation to ensure fair assessment practice. The results from this study prove that it is not possible to use findings from other cross cultural studies on the test, as support for a rationale to use the test on a local population. Sample characteristics and context play critical roles in performance on tests.

The study's findings have emphasised the importance of using appropriate local normative data. The results show how damaging it can be to use inappropriate norms to interpret performance in a different cultural population. The use of these norms and the resulting interpretations, have far reaching consequences, some of them extremely unethical. False diagnosis in terms of learning disorder, brain injury or mental retardation may be made, which has long-term ramifications for the child and family. This may lead to inappropriate distribution of state resources. Cognitive assessment and its influence in determining access to resources implies that for many children

who are inaccurately identified through the misinterpretation of test scores future opportunities are out of reach.

Moreover, the danger in making decisions based only on the results of cognitive test performance is clear. Performance on a test of cognitive ability should not be used on its own to indicate an individual's true ability. Vygotsky would view this as only an indication of the past opportunities that the individual had in learning specific skills. In fact, Vygotsky's idea of focusing on learning potential takes into account the dynamic nature of cognitive ability and the impact that experience or environmental factors play in the development of specific skills (Bedell et al., 1999). It is widely recognised that test performance must be considered in a context which covers various aspects of an individual's psychological and physical functioning. This is especially important in a setting like South Africa where many children are underprivileged and are at risk for harmful factors associated with poverty (Luiz, 1994, in Luis & Jansen, 2001). In fact Elkonin, Foxcroft, Roodt and Astbury (2001) state that "the use of a single test score to make decisions regarding the child should be seen as an unethical assessment practice" (p.278). Considering that the present study has been unable to make conclusive reports on the validity of the RCPM for this sample, practitioners should use the test with caution and as part of a larger assessment battery, when assessing Zulu-speakers.

It is obvious that more extensive research is required in the area of multicultural testing and the development of appropriate local norms must be seen as urgent.

5.6 RECOMMENDATIONS FOR FURTHER RESEARCH

Further research using the RCPM should focus on remedying some of the flaws already highlighted in the design of the present study. More context variables such as SES and level of parental education need to be included, as do better operationalisations of nutrition and acculturation. Possible sources of bias, which may threaten the results are better identified before the start and incorporated into the design. This facilitates explanations of observed differences using measured effects, as opposed to speculation of unknown factors and their possible role. "In an ideal study the set of context variables will be chosen in such a way that the remaining effect for culture will be zero" (Poortinga & Van de Vijver, 1987, p.271).

Although the reliability of the RCPM for this sample is high, the validity of the test remains uncertain. Construct validity thus requires further investigation. This should include DIF analysis based on a model which Raven decides is appropriate.

It would be interesting to conduct retesting in future studies, in order to examine the theory which Wober (1969) offers to account for practice gains. This could be carried out with the use of both the RCPM and the parallel version of the test. The effect of familiarity with test content and testing situation is likely to be more marked with such a comparison. Retest reliability could also be established.

Considering the difficulty with an unfamiliar test format and the research which has shown different perceptual skills for African samples, the board form of the test may be more appropriate than the book form. This would allow for physical manipulation of the puzzles, thereby encouraging direct learning experience. It is suggested that practice examples using a piece of brightly patterned African fabric with a piece missing, might also facilitate a more familiar medium of test presentation.

As shown in the review of current literature, little research has been conducted with the RCPM in South Africa. Furthermore, much of this research has been undertaken by master's students. It

seems that the RCPM is a test worthy of larger scale evaluation and investigation. It is a test which is widely used in South Africa, particularly favoured because of the language barrier which exists in many of the settings, where it is used by psychologists. It is also a relatively quick test to administer and cost effective. Both these factors add to its popularity. Research into the RCPM by a larger organisation would allow for more controlled studies, larger, more representative samples and access to better expertise in the field. It would also facilitate a central database, where differential, local normative data may be collected and published. Comparison of performances across the different cultural groups within South Africa should be made. Given this study's results, which show the significant effect of culture and different opportunities for experience and education, it is likely that performances across these groups will vary. This would indicate that local norms will need to be organised into differential norms.

Funding in the area of test evaluation and adaptation must receive greater priority. It is clear that at present many test users may be using inappropriate tests and/or norms. The ethical implications of this have been noted.

5.7 SUMMARY AND CONCLUSIONS

It has been shown that cultural equivalence is essential in any comparison that is being made between two different groups of people (Geisinger, 1994; Helms, 1992; Helms-Lorenz & Van de Vijver, 1995; Poortinga, 1987; Poortinga, 1995; Poortinga & Van de Vijver, 1987; Scheuneman, 1979). This involves establishing whether a test is appropriate for each group, which may be shown if no bias exists on different levels of the test design and use. In South Africa, research into test adaptation and the development of local norms has been insufficient (Anderson, 2001; Shanahan et al., 2001; Shuttleworth-Jordan, 1996; Viljoen et al., 1994). This study aimed to address this shortfall by investigating a commonly used assessment measure, the RCPM.

The findings of the study confirm the importance of determining a test's suitability before using it in a new population. Investigating the psychometric properties of the test for this sample, revealed good reliability and highlighted possible sources of bias at the item level. Some items appeared to out of sequence in the ordering of difficulty and some seemed to have poor discriminatory power for this sample. The correct analysis procedures for DIF were not performed, which limits the irrevocability of these findings. These findings suggest that construct validity can not be verified and that further research in this area must be undertaken. Results which did show positive evidence for construct validity, were the reliability coefficient and factor structure similar to those reported in other studies.

The other aspect of this study focused on a comparison between the study's sample performances with performances of the British standardisation sample. These particular norms were selected for comparison because they are commonly used for interpretation in South Africa. The results revealed a significant difference between the two sample's performances. The Zulu sample's scores were on average, 7.5 points below those on the British sample. These findings suggest that the British norms are inappropriate for this sample. Furthermore the pattern of differences observed, where differences increased with age, challenges some of the theoretical underpinnings of the test's design. It is suggested here that Vygotsky's theory of cognitive development provides a better explanation of the Zulu sample's performance. It was further

suggested that the lower performance of Zulu children should not be interpreted as lower intelligence, but can be explained by many contextual factors. Culture is considered a strong moderating variable and tends to impact on other factors like perceptual skills and education. Nutrition, SES and parental education are also indicated as contributing factors. Due to limitations in the design of the study, much of this discussion remained speculation rather than fact.

The relationship between sample characteristics was also explored. Performance and age were significantly correlated. Level of education however, appeared to be the strongest predictor of performance on the RCPM for this sample. These findings have not been reported elsewhere. Gender and location both play a significant role in performance. Improved performance for males was observed and urban children performed considerably better than rural children. These results again highlight the importance of several contextual factors. Education and experience are considered particularly important, while acculturation appears to have a moderating influence on performance.

The following conclusions may be drawn from the present study: The British normative data for the RCPM is not appropriate for interpretation of Zulu-speaking children's scores. The appropriateness of the RCPM for this sample is inconclusive. Results did not satisfy all criteria for construct validity. Some of the analysis however, requires further investigation. If construct validity can be shown, it would still appear that the RCPM requires adaptation for use with a Zulu speaking sample. Developing local normative data for this sample must be considered a critical priority for further research.

REFERENCES

Anastasi, A. (1990). *Psychological Testing.* New York: Macmillan.

Anderson, S.J. (2001). On the importance of collecting local neuropsychological normative data. *South African Journal of Psychology, 31*(3), 29-34.

Ardila, A. (1995). Directions of research in cross-cultural neuropsychology. *Journal of Clinical and Experimental Neuropsychology, 17*(1), 143-150.

Ardila, A., & Rosselli, M. (1994). Development of Language, Memory and visuospatial Abilities in 5-12-year-old children using a neuropsychological battery. *Developmental Neuropsychology, 10*(2), 97-120.

Ardila, A., Rosselli, M. & Rosas, P. (1989). Neuropsychological Assessment in Illiterates: Visuospatial and Memory abilities. *Brain and Cognition, 11*, 147-166.

Baron, R.M. & Kenny, D.A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic and statistical considerations. *Journal of Personality and social psychology, 51*(6), 1173-1182.

Bass, N.T. (2000). The Raven's Coloured Progressive Matrices test: A pilot study for the establishment of normative data for Xhosa-speaking primary school pupils in the Grahamstown region. Rhodes University; Department of Psychology. Unpublished Masters thesis.

Bedell, B., Van Eeden, R., Van Staden, F. (1999). Culture as moderator variable in psychological test performance: Issues and trends in South Africa. *South African Journal of Industrial Psychology, 25*(3), 1-7.

Berk, L.E. (2000). *Child development.* Boston: Allyn and Bacon.

Berry, J.W. (1971). Ecological and cultural factors in spatial perceptual development. *Canadian Journal of Behavioural Science, 3*, 324-336.

Berry, J.W. (1997). Preface. In J.W. Berry, Y.H. Poortinga & J. Pandey (Eds.), *Handbook of Cross-Cultural Psychology: Vol. 1. Theory and Method* (pp. x-xv). Boston: Allyn & Bacon.

Berry, J.W., Poortinga, Y.H., Dasen, P.R., & Segall, M.H. (Eds.). (1992). *Cross-cultural psychology: Research and applications*. Cambridge: Cambridge University Press.

Brislin, R. (1980). Translation and content analysis of oral and written material. In H.C. Triandis & J.W. Berry (Eds.), *Handbook of Cross-Cultural Psychology: Volume 2. Methodology* (pp.389-444). Boston, Massachusetts: Allyn and Bacon Inc.

Carlson, J.S. (1973). A note on the relationship between the Raven's Coloured Progressive Matrices and operational thought. *Psychology-in-the-Schools, 10*(2), 211-214.

Carlson, J.S. & Jensen, M.J. (1980). The factorial structure of the Raven Coloured Progressive Matrices test: A reanalysis. *Educational and Psychological Measurement, 40*, 1111-1116.

Carlson, J.S. & Jensen, M.J. (1981). Reliability of the Raven Coloured Progressive Matrices test: Age and ethnic group comparisons. *Journal of Consulting and Clinical psychology, 49*(3), 320-322.

Carpenter, P.A., Just, M.A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices test. *Psychological Review, 97*(3), 404-431.

Centres for Disease Control and Prevention (2005). BMI-Body Mass Index: BMI for children and teens. Retrieved November 15, 2005, from http://www.cdc.gov/nccdphp/dnpa/bmi/bmi-for-age.htm

Costenbader, V., & Ngari, S.M. (2001). A Kenya standardization of the Raven's Coloured Progressive Matrices. *School Psychology International, 22*(3), 258-268.

Court, J. H. & Raven, J. (1995). *Manual for Raven's Progressive Matrices and Vocabulary scales: Section 7 – Research and References.* Oxford: Oxford psychologists Press Ltd.

Crawford-Nutt, D.H. (1976). Are black scores on Raven's Standard Progressive Matrices an artifact of method test presentation? *Psychologia Africana, 16*, 201-206.

Crawford-Nutt, D.H. (1977). The effect of educational level on the test scores of people in South Africa. *Psychologia Africana, 17*, 49-59.

Das, J.P., Kirby, J.R., & Jarman, R.F. (1975). Simultaneous and successive synthesis: An alternative model for cognitive abilities. *Psychological bulletin, 83*(1), 87-103.

Durrheim, K. (2005). The standard normal distribution. In C. Tredoux & K. Durrheim (Eds.), *Numbers, hypotheses & Conclusions: A course in Statistics for the social sciences* (pp.90-107). Lansdowne: UCT Press.

Elkonin, D., Foxcroft, C., Roodt, G. & Astbury, G. (2001). The use of assessment measures in various applied contexts. In C. Foxcroft & G. Roodt (Eds.), *An introduction to psychological assessment in the South African context* (pp. 265-294). Oxford: Oxford University Press.

Emerling, F. (1990). An investigation of test bias in two nonverbal cognitive measures for two ethnic groups. *Journal of Psychoeducational Assessment, 8*, 34-41.

Finchilescu, G. (2005). Measurements. In C. Tredoux & K. Durrheim (Eds.), *Numbers, hypotheses & Conclusions: A course in Statistics for the social sciences* (pp.201-229). Lansdowne: UCT Press.

Fletcher, J.M., Todd, J. & Satz, P. (1975). Culture-fairness of three intelligence tests and a short-form procedure. *Psychological Reports, 37*, 1255-1262.

Foxcroft, C. (2001). Developing a psychological measure. In C. Foxcroft & G. Roodt (Eds.), *An introduction to psychological assessment in the South African context* (pp. 69-102). Oxford: Oxford University Press.

Foxcroft, C., Roodt, G. & Abrahams, F. (2001). The practice of psychological assessment: Controlling the use of measures, competing values, and ethical practice standards. In C. Foxcroft & G. Roodt (Eds.), *An introduction to psychological assessment in the South African context* (pp. 105-128). Oxford: Oxford University Press.

Geisinger, K.F. (1994). Cross-cultural normative assessment: translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological assessment, 6*(4), 304-312.

Hammond, S. (1995). Introduction to Multivariate Data analysis. In G.M. Breakwell, S. Hammond & C. Fife-Schaw (Eds.), *Research Methods in psychology* (pp.360-385). London: Sage Publications Ltd.

Helms, J.E. (1992). Why is there no study of cultural equivalence in standardized cognitive ability testing? *American Psychologist, 47*, 1083-1101.

Helms-Lorenz, M. & Van de Vijver, F.J.R. (1995). Cognitive assessment in education in a multicultural society. European Journal of Psychological assessment, 11(3), 158-169.

Hudson, W. (1962a). Cultural problems in pictorial perception. *South African Journal of Science, 58*, 189-196.

Hudson, W. (1962b). Pictorial Perception and educational adaptation in Africa. *Psychologia Africana, 9*, 226-239.

Jahoda, G. & Krewer, B. (1997). In J.W. Berry, Y.H. Poortinga & J. Pandey (Eds.), *Handbook of Cross-Cultural Psychology: Vol. 1. Theory and Method* (pp.1-42). Boston: Allyn & Bacon.

Jinabhai, C.C., Taylor, M., Rangongo, M.F., Mkhize, N.J., Anderson, S., Pillay, B.J. & Sullivan, K.R. (2004). Investigating the mental abilities of rural Zulu primary school children in South Africa. *Ethnicity and health, 9*(1), 17-36.

Kanjee, A. (2001). Cross-cultural test adaptation and translation. In C. Foxcroft & G. Roodt (Eds.), *An introduction to psychological assessment in the South African context* (pp. 86-102). Oxford: Oxford University Press.

Kaufman, C., Maharaj, P., & Richter, L. (2001). Children's schooling in South Africa: Access, quality and geography. In Equity and Diversity in Education: Commonwealth snapshots. Pre-Oxford Conference (pp. 32-45). Oxford: Acorn Publishing.

Kendall, I.M. (1976). The Predictive validity of a possible alternative to the classification test battery. *Psychologia Africana, 16*(2), 131-146.

Kimura, D. (2000). *Sex and cognition.* Cambridge, Massachusetts: Bradford Book, MIT Press.

Kirby, J.R., & Das, J.P. (1978). Information processing and human abilities. *Journal of Educational Psychology, 70*(1), 58-66.

Klippel, M.D. (1975). Measurement of intelligence among three New Zealand ethnic groups. *Journal of Cross-cultural psychology, 6*(3), 365-376.

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

Owen, K. (1992). The suitability of Raven's Standard Progressive Matrices for various groups in South Africa. *Personality and Individual Differences, 13*(2), 149-159.

Owen, K. (1998). The role of psychological tests in education in South Africa: Issues, controversies and benefits. Pretoria: Human Sciences Research Council.

Poortinga, Y.H. (1989). Equivalence of cross-cultural data: an overview of basic issues. *International Journal of Psychology, 24*, 737-756.

Poortinga, Y.H. (1995). Cultural bias in assessment: Historical and thematic issues. *European Journal of Psychological Assessment, 11*(3), 140-146.

Poortinga, Y.H. & Van de Vijver, F.J.R. (1987). Explaining cross-cultural differences: Bias analysis and beyond. *Journal of Cross-cultural psychology, 18*, 259-282.

Raven, J.C. (1965). *Guide to using the Coloured Progressive Matrices*. London: H. K. Lewis & Co. Ltd

Raven, J., Raven, J.C., & Court, J.H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales: General Overview 1998 Edition: Section 1*. San Antonio, TX: Psychological Corporation.

Reddy, P., Knowles, A. & Reddy, S. (1995). Language issues in cross-cultural testing. *Australian Psychologist, 30(1), 27-29.*

Rosnow, R.L. & Rosenthal, R. (1996). *Beginning Behavioral Research*. New Jersey: Prentice-Hall Inc.

Scheuneman, J. (1979). A method of assessing bias in test items. *Journal of Educational Measurement, 16*(3), 143-152.

Schmidtke, A. & Schaller, S. (1980). Comparative study of factor structure of Raven's Coloured Progressive Matrices. *Perceptual and Motor skills, 51*, 1244-1246.

Serpell, R. (1979). How specific are perceptual skills? A cross-cultural study of pattern reproduction. *British Journal of Psychology, 70*, 365-380

.

Shanahan, S.F., Anderson, S.J., & Mkhize, N.J. (2001). Assessing psychological distress in Zulu-speakers: preliminary findings from an adaptation of the SCL-90-R. *South African Journal of Psychology, 31*(4), 1-12.

Shuttleworth-Jordan, A.B. (1996). On not reinventing the wheel: A clinical perspective on culturally relevant test usage in South Africa. *South African Journal of Psychology, 26*(2), 96-102.

Sigmon, S. (1983). Performance of American schoolchildren on Raven's Coloured Progressive Matrices scale. *Perceptual and motor skills. 56*, 484-486.

Skuy, M., Gewer, A., Osrin, Y., Khunou, D., Fridjhon, P., & Rushton, P. (2002). Effects of mediated learning experience on Raven's matrices scores of African and non-African university students in South Africa. *Intelligence, 30*, 221-232.

South Africa Online Travel Guide (2005). *Pietermaritzburg, KwaZulu-Natal.* Retrieved December 14, 2005, from http://www.southafrica-travel.net/kwazulu/e2pietmb.htm

Statistics South Africa (2005). South African Statistics, 2004/05.

Statistics South Africa (2005). Population census, 2001.

Terre Blanche, M. (2005). Central Tendency. In C. Tredoux & K. Durrheim (Eds.), *Numbers, hypotheses & conclusions: A course in statistics for the social sciences* (pp.201-229). Lansdowne: UCT Press.

Thomas, R.M. (1985). *Comparing theories of child development*. California: Wadsworth Publishing company.

Van de Vijver, F.J.R. & Leung, K. (1997). Methods and data analysis of comparative research. In J.W. Berry, J. Pandey & Y.H. Poortinga (Eds.), *Handbook of Cross-cultural Psychology: Volume 1 Theory and Method* (pp.257-300). Boston: Allyn and Bacon.

van Eeden, R. & de Beer, M.(2001). Assessment of cognitive functioning. In C. Foxcroft & G. Roodt (Eds.), *An introduction to psychological assessment in the South African context* (pp. 175-193). Oxford: Oxford University Press.

Vass, V.A. (1992). Standardization of Raven's Standard Progressive Matrices for secondary school African pupils in the Grahamstown region. Unpublished master's thesis, Rhodes University, Grahamstown.

Viljoen, G., Levett, A., Tredoux, C., Anderson, S. (1994). Using the Bender Gestalt in South Africa: Some normative data for Zulu-speaking children. *South African Journal of Psychology, 24*(3), 145-151.

Vygotsky, L.S. (1978). Mind in society: The development of higher psychological processes. Cambridge, Mass.: Harvard University Press.

Whaley, S.E., Sigman, M., Neumann, C., Bwibo, N., Guthrie, D., Weiss, R.E., Alber, S. & Murphy, S.P. (2003). The impact of dietary intervention on the cognitive development of Kenyan school children. *The Journal of Nutrition, 133*(11), 3965-3971.

White, D. & Panunto, B. (1978). Verbal and nonverbal abilities in English first and second language children. *Psychological Reports, 42*, 191-197.

Wiedl, K.H., & Carlson, J.S. (1976). The factorial structure of the Raven Coloured Progressive Matrices test. *Educational and Psychological Measurement, 36*, 409-413.

Wober, M. (1969). The meaning and stability of Raven's Matrices test among Africans. *International journal of psychology, 4*(3), 229-235.

Wolfaardt, J.B. (2001). Basic concepts. In C. Foxcroft & G. Roodt (Eds.), *An introduction to psychological assessment in the South African context* (pp. 34-68). Oxford: Oxford University Press.

Zindi, F. (1994). Differences in psychometric performance. *The Psychologist, 7*, 549-552.

APPENDIX A

School of Psychology
P/Bag X01 Scottsville
PIETERMARITZBURG, 3209
South Africa
Phone: +27 33 2605853
Fax: +27 33 2605809

UNIVERSITY OF
KWAZULU-NATAL

CONSENT FORM FOR PARENTS/GUARDIANS

Dear Parent / Guardian,

Your child has been invited to be part of a research study run by a researcher from the University of KwaZulu Natal, School of Psychology, with the cooperation of the Education Department of Natal.

The study involves a test called the Raven's Coloured Progressive Matrices. It is often used in schools to measure a child's ability and in hospitals, it is used to see if there are any problems as a result of brain damage. This test was made overseas and so the scores, which we have to compare our children to, are from children overseas. This causes problems because sometimes our children will look like they are doing poorly when they are not, or could look like they are doing well when there could be some problems. We need to compare our children to scores that have been collected from our own population. It is very important to collect this information.

This study is to do just that. The researcher would like to have your child help to collect this information. Your child will be given the test to do. It will be done at school in the classroom with the other pupils and the teacher. It is a short test, only taking about 20 to 30 minutes and just involves looking at pictures and picking out the missing pieces. The child will also have to fill out a questionnaire that gives details about their name, age, weight, height and grade at school. The researcher will measure things like weight and height. Your child's name and participation

in this study will be kept confidential as well as the name of the school. If the study is published, identifying details will be protected very carefully.

Your child will also benefit from doing the test. The results will only be available to the parents or guardians. The results will give you information about your child's ability and will also let you know if there are any problems that need working on. If some problems are identified then the researcher will notify you and ask you to get in touch with the school, which will be able to refer your child for appropriate help.

Your child will also be asked whether they want to participate in the study and will be free to refuse to do it. Even if they agree and later change their minds, they can also still refuse. You are also free to withdraw your consent at any time.

.....

THE APPROPRIATENESS OF THE RAVEN'S COLOURED PROGRESSIVE MATRICES FOR ZULU SPEAKING CHILDREN

I hereby give consent for

Name of child:_____

Date of birth:_____

School grade currently completing:_____ to participate in this study.

_____ _____

Signature of parent/guardian Date

We appreciate your understanding in this matter and thank you for you cooperation.

Regards

Tarryn Kihn
(Masters Psychology Student)

APPENDIX B

School of Psychology
P/Bag X01 Scottsville
PIETERMARITZBURG, 3209
South Africa
Phone: +27 33 2605853
Fax: +27 33 2605809

UNIVERSITY OF
KWAZULU-NATAL

INCWADI YESICELO SEMVUME YABAZALI

Ingane yakho iyamenywa ukuba ingxenye yocweningo eyenziwa abacwaningi base nguvesi yekwaZulu/Natal noMnyengo wezeMfundo kuleNyuvesi.

Lolu cwaningo lumayelana nesivivinyo esibizwa npokuthi Ravens Coloured Matrices. Lesivivinyo siveme ukusetshezìswa eziikoleni ukuhlola ikhono labantwana kanjelo nasezibhedlela nkubheka ukuphazemiseka komqondo. Lesivivinyo senziwa phesheya kenja nemamaki, lokhu-ke kube sekughatheniswa nabantwana bethu kuleli. Lokhu-ke kuye kudale inkinga yokuthi kwesinye isikhethi abantwana bethu kuleli beyeye bathole ngephansi kwalabo baphesheya, noma kube ngethi benza kancono kunebaphesheya yize bengeba nezinkinga ezicoshile. Sidinga-ke ukuphathanisa imiphumela yabantwana bethu kuleyo esivele ikhona. Ngekho-ke kubelulkile ukugogo lolulwazi.

Lolucweningo olokwenza lokho. Umcwaningi angethanda ukucela ingane yakho imsize ukugoga loku lwezi. Lesivivingo siyonikeswa ingane yakho, esikoleni kanye nobanye abantwana nothisha webo. Lesi isivivinyo esitushane, sithatha ushe imizuzu engu zo kuya kwewu 30, kenti simayelena nokugcwelisa isithombe ngenxenye eshodaya. Ingane kuyomele igcwelise igama layo, iminyeka, isisindo sayo, ubude beyo kanye neklasi ekulona esikoleni. Umcweningi uzoyikala isisìndo nobude beyo ingane. Igama lengane nesikole kuyogodlwa, uma sekushicilelwa imiphumela yalolukhuaningo.

Ingane yakho izozuza kulesivivinyo. Imiphumela inganikezwa abezali ume beyifuna, futhi inganikezwa isikole noma othisha uma umzali evvma. Imiphumela iyochaza ngekhono lengane yakho futhi nanokusho ukuthi inozo yini isinkinga ezidingwa ukuqeshelwa. Uma zikhona izinkinga uyokwazisa mzali, bese uxhumena nesikole somntwana ukuze asizekale umntana.

Ngaphombi kokwenza ucweningo, umntwana wakho uyobuzwa ukushi uya thanda yini ukuba yingxenye yophenyo ale noma avvume. Uma evvuma abuye ashintshe ale wamukelekile ukuhoxisa imvvme yekho.

.....

UCWANINGO LWESIVIVINYO SE RAVENS COLOURED PROGRESSIVE MATRICES KUBANTWANA ABAKHULUMA ISIZULU

Ngiyavuma ukushi (Igama lamntwana) u_____

ofunda uklasi _____ yelolu cwaningo. Usuku lomntwana lokuzalwa

ngu:_____

_____ _____

Signature Usuku

Siyabonga ukusixhasa kuvakho kulolu cwaningo.

Tarryn Kihn
(isitshudent se masters Psychology)

APPENDIX C

PROVINCE OF KWAZULU-NATAL
ISIFUNDAZWE SAKWAZULU-NATAL
PROVINSIE KWAZULU-NATAL

DEPARTMENT OF EDUCATION AND CULTURE
UMNYANGO WEMFUNDO NAMASIKO
DEPARTEMENT VAN ONDERWYS EN KULTUUR

Address: 228 Pietermaritz Street
Pietermaritzburg
3201

Private Bag
Pietermaritzburg
3200

Telephone: 033-3425481
Fax: 033-3425490

Enquiries: Dr DNW NGCOBO
Reference:
Date: 21/09/2004

Ms Tarryn Kihn
School of Psychology
University of KwaZulu-Natal

FAX: 033-2605809

RE: PERMISSION TO CONDUCT RESEARCH AT SCHOOLS IN PIETERMARITZBURG

Please be advised that your request has been forwarded to Mr BH Mthabela, the Director of Research Strategy & EMIS.

For further correspondence kindly contact Mr Mthabela at the following contact numbers:

Telephone: 031-274 4924/26
Fax: 031-274 4922

Looking forward to the results of your research.

ACTING REGIONAL CHIEF DIRECTOR

APPENDIX D

RCPM Test Instructions (Translated into Zulu)

Bheka lokhu.

Uyabona kulephethini kunesinqamu esikhishiwe. Kuzona zonke lezi zinqamu ezingezansi...

Siyi esifenele ukugcwalisa isikhala, kodwa sinye kuphela esiyiphethini elifanele. Unamba 1 uyisimo esifanele, kodwa akulona iphethini elifanele. Unamba 2 akulona iphethini elifanele okwempela. Unamba 3 akalungile. Unamba 6 ucishe abe ofanele, kodwa akalungile la.

Kukodwa okulungile. Khomba isinqamu esifanele ukuqedela iphethini.

Manje khomba isinqamu esingena la.

Bhekisisa kulezi zinqamu.

Sisodwa kuphele isinqamu esifanele ukuqedela iphethini. Qaphela. Bheka ngesinye kulezi zinqamu eziyisithupha kuphela.

Manje khomba esifanele ukungena la.

Yibeso yini esifanele ukungena la?

Kulungile, khomba leso esifanele.

Yileso yini esifanele?

Yisiphi ocabanga ukuthi yisona sona?
Bheka iphethini ngokucophelele. Manje isiphi kulezi zinqamu...

Esingena la?

Qaphelisisa, sisodwa esifanele. Yisiphi sona? Qiniseka ngokuthola esifanele ngophembi kokusikhombe.

Uyabona ukuthi kwenziwa kanjani? Lokhu. Lokhu. Lokhu. Kuzoba yisiphi la? Khomba esifanele kulezi esizongena la. Qaphela. Bhekisisa ngosinye. Sisodwa okuyisona lana. Yisiphi leso?

Yileso yini esifanele ukuqedela iphethini?

Bheka ngokucophelela iphethini.
Qaphela. Sisodwa kuphela kulezi zinqamu esiqedela iphethini ngokufanele.

Yisiphi?

Ungayeki incwejana ye-test uze ucelwe ukwenza kanjalo.

Chaza ukuthi iProgressive Matrices Test ikala ikhona lokucabanga ngokucacileyo.
Vula incwajana yakho yeCPM inkinga yokuqala (A1)
Kukanje. Phezulu kuthi "A" lokhu kusho I sethi A ... futhi kwehlukanisiwe la
Ephepheni lakho lezimpendulo ka sethi A

Lena inkinga uA1. Uyabona ukuthi iyini. Ngenhla iphethini elinento ekhishiwe. Enye nenye yalezi zinqamu ngezansi.

Iyisimo esifanele ukugcwalisa lesi khala kodwa-ke ukuze zonke ziqedele leli phethini. Unamba 1...

Uyisimo esifanele kodwa akalona iphethini esifanele unamba 2 no 3...

Abalungile. Bayagcwalisa indawo kodwa abalona iphethini elifanele. Eqinisweni unamba 2 akalona iphethini ngempela. No namba 3 akalungile. Unjani unamba 6? Uyiphethini esifanele?

Kodwa akagcwele yonke indawo ngakho angashiye isikhala. Beka iminwe yakho kuleso akayisona lana.

Yebo, unamba 4 uyena yena. Ngakho impendulo ka A1 ngu "4". Bheka u "4" ekuqaleni kuka namba 1 ngokwesikhala sokuqala. Siyacela ungamaki incwajana ye test. Ungephenyi ngale okwamanje.

Kuwo wonke anakhasi alencwajana kunesinqamu esikhishiwe. Kufanele ukhethe ukuthi kulezi zinqemu ngezansi yisiphi esigeson esalephethini uma ucabaga ukuthi ususitholile isinqemu esifanele, bhala inamba yaso kwinamba efanele ephepheni lempendulo.

Ume wenza iphutha noma ufuna ukushintsha impendulo yakho, yenza isiphambano empendulweni engafanele bese ubhala inamba yempendulo efanele.

Ungayihlikihli impendulo engafanele ungabheki kwincwajana ye test. Izinkinga zilula ekuqeleni ziye ziba nzima ume uqhubeka.

Awuphicwaphicwa. Ume ubhekisisa indlela izimpendulo ezitholakala ngayo, uzothola ukuthi uma uqhubeka ngeke kube nzima kakhulu.

Zama ngokunye, kusukela ekuqaleni kuze kube sekugcineni kwencwajana ukuqopha izimpendulo ephepheni lezimpendulo. Uma uqopha izimpendulo qiniseka ukuthi uqopha esikhaleni esifanele kanye naseduze kwenamba efanele.

Sebenza ngokwamandla akho.

Ungashiyi lutho. Uma ungaqinisekanga, qaphela, ngoba kwesinye isikhathi kube okulungile. Uma uhluleka dlulela phambili bese ubuya kukho kamuva.

Thatha usikhathi sakho.
Phenya uzame inkinga ewu A2

Impendulo efanele ngu namba 5. Senibhale nonke u "5" kunamba 2 esikhaleni sokuqala sephepha lezimpendulo?

Qzoqubeka kanjena ngokwakho uze ufike ekugcineni kwencwajana. Ngizolokhu ngizobheka ukuthi usaqhuba kahle yini.
Unayo imibuzo?

Khumbula, ungacishi amaphutha. Beka isiphambano empendulweni engalungile bese ubhala inamba efanele.

Phenya uye enkingeni ewu A3 bese uqala.

APPENDIX E

Table 1

Description of sample by age

Age	n	Gender		Age group category
		Male	Female	
5	2	2	0	4.09-5.02
5.5	3	1	2	5.03-5.08
6	24	15	9	5.09-6.02
6.5	23	9	14	6.03-6.08
7	32	13	19	6.09-7.02
7.5	40	20	20	7.03-7.08
8	46	27	19	7.09-8.02
8.5	53	27	26	8.03-8.08
9	64	34	30	8.09-9.02
9.5	54	24	30	9.03-9.08
10	65	32	33	9.09-10.02
10.5	50	25	25	10.03-10.08
11	33	16	17	10.09-11.02
11.5	22	14	8	11.03-11.08
12	8	3	5	11.09-12.02
12.5	2	0	2	12.03-12.08
13	1	1	0	12.09-13.02

Table 2

Post-hoc comparisons of mean differences in scores across grades

		Grade (J)	Mean Difference I-J	Sig (*) .05
Grade (I)	1	2	-0.44	.53
		3	-2.36*	.001
		4	-4.17*	.000
		5	-4.83*	.000
	2	1	0.44	.529
		3	-1.92*	.006
		4	-3.73*	.000
		5	-4.39*	.000
	3	1	2.36*	.001
		2	1.92*	.006
		4	-1.82*	.012
		5	-2.48*	.010
	4	1	4.17*	.000
		2	3.73*	.000
		3	1.82*	.012
		5	-0.66	.509
	5	1	4.83*	.000
		2	4.39*	.000
		3	2.48*	.010
		4	0.66	.509

* Pairwise significances as determined by Levene's test

Table 3

Components Extracted and Eigenvalues

ponent	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
	6.378	17.716	17.716	6.378	17.716	17.716	5.100	14.168	14.168
	2.702	7.506	25.223	2.702	7.506	25.223	3.727	10.353	24.521
	1.977	5.493	30.715	1.977	5.493	30.715	2.230	6.194	30.715
	1.611	4.475	35.190						
	1.270	3.528	38.718						
	1.183	3.286	42.004						
	1.139	3.164	45.168						
	1.109	3.081	48.249						
	1.055	2.930	51.179						
	1.047	2.908	54.087						
	.948	2.633	56.720						
	.929	2.580	59.301						
	.898	2.494	61.795						
	.871	2.418	64.213						
	.852	2.366	66.579						
	.803	2.231	68.810						
	.776	2.156	70.965						
	.768	2.134	73.099						
	.709	1.970	75.070						
	.694	1.929	76.999						
	.680	1.890	78.889						
	.655	1.819	80.707						
	.635	1.764	82.472						
	.611	1.698	84.170						
	.597	1.659	85.829						
	.578	1.605	87.434						
	.550	1.528	88.962						
	.523	1.452	90.414						
	.508	1.410	91.824						
	.456	1.266	93.090						
	.448	1.246	94.335						
	.437	1.214	95.550						
	.428	1.188	96.738						
	.419	1.163	97.901						
	.406	1.129	99.030						
	.349	.970	100.000						

Extraction Method: Principal Component Analysis.

Table 4

Principal component analysis with oblique rotation: Pattern matrix

Item	Component		
	1	2	3
A1	-.055	**-.401**	.076
A2	-.086	**-.468**	.068
A3	-.039	**-.495**	-.002
A4	.066	**-.434**	-.050
A5	.117	**-.417**	.040
A6	.125	**-.472**	-.001
A7	**.400**	-.002	.033
A8	.142	-.110	-.053
A9	.243	-.132	.051
A10	.340	-.132	.027
A11	.171	.004	-.003
A12	.327	.184	-.028
Ab1	.018	**-.609**	-.013
Ab2	.144	**-.496**	.019
Ab3	.036	**-.562**	-.041
Ab4	**.586**	-.087	-.008
Ab5	**.588**	.031	-.029
Ab6	**.610**	.019	**.068**
Ab7	**.644**	-.005	**.068**
Ab8	**.692**	.020	**.015**
Ab9	**.436**	.074	**.095**
Ab10	**.423**	-.046	.034
Ab11	.280	-.056	-.025
Ab12	.034	.184	-.009
B1	.182	**-.485**	-.110
B2	.285	**-.433**	-1.27
B3	**.514**	-.207	.010
B4	**.540**	-.117	-.007
B5	**.579**	.006	**.078**
B6	.397	-.067	.109
B7	.241	.035	.227
B8	.096	.033	**.604**
B9	.058	-.027	**.550**
B10	.001	-.087	**.534**
B11	.028	-.026	**.650**
B12	.138	-.031	-.091

Note: loadings ≥ .40 considered relevant for interpretation of factors