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.

Natalie Bass

Submitted in partial fulfillment of the requirements for the degree of Master of Social Science. (Clin Psych) in the Department of Psychology, Rhodes University, Grahamstown

January 2000

Supervisor: Jannie Knoetze

Abstract

The Raven's Coloured Progressive Matrices (CPM) test is used extensively across a wide variety of settings in South Africa, however more appropriate local normative data has yet to be established. The CPM is internationally recognised as a culture-fair test of nonverbal intelligence, designed for use with children between the ages of 5½ and 11½. This pilot study thus sought to establish normative data for this instrument for a population of Xhosa-speaking Primary School children in the peri-urban township area in the Grahamstown region. The booklet version of the test was used and it was administered in group format and according to an alternate method of test administration (using Xhosa instructions) developed by Vass in 1992. The final normative sample consisted of 197 male and 182 female Xhosa-speaking children in Grades Two to Seven (N=379).

The results showed (1) a significant effect of age on test scores, where scores increased with age as expected; (2) a consistent tendency for males to outperform females was also noted, however small sample sizes precluded any categorical claims to this effect; (3) no significant effect of education on test scores was observed and finally; (4) and finally, it appeared that the norms generated for this study revealed a tendency to be lower than those obtained by Raven, Court and Raven (1990) during the standardisation of this instrument in the United Kingdom and America. The study concluded that (1) there is an urgent need for more appropriate South African normative data for this test; and (2) that when assessing African children from disadvantaged backgrounds, further research into the effects of cultural and socio-economic factors and gender on non-verbal intelligence (and on performance on this test in particular) is required.

Declaration

The author hereby declares that this whole thesis, unless specifically indicated to the contrary in the text, is her own work.

_____

Natalie Bass

Acknowledgements

Jannie, thanks for being there, especially in the very beginning and at the end…..

I would also like to take this opportunity to thank the staff at the Department of Education for their cooperation and assistance on this project; and the children and staff at the Primary school involved, for their enthusiastic participation and the use of their facilities.

My deepest thanks also to Thabile Hlengane of the Grahamstown Education Support Centre; for sharing her vast experience as a Remedial teacher working in schools in this area with me, and for her extensive contributions as a test administrator, interpreter and field worker.

In addition, this study would not have been possible without the funding supplied by the C.S.D. (Centre for Scientific Development) as well as the generosity of the test's developer - Mr. John Raven who (when informed of the lack of availability of test booklets in the Eastern Cape) offered 45 test booklets on loan. Your support is gratefully acknowledged.

On a more personal note:

- to Brendan – thanks especially for the very necessary post-M2 trip to Mozambique, where you and Dois-M made it all seem worthwhile…..but more importantly, thank you for accompanying me on this journey for the past six years - you being there for me truly made all the difference.
- to Renate – thanks for all the help, endless patience, your outstanding ability to 'work with large large documents'and for making the whole process of editing and printing a master's thesis virtually stress-free.
- to my father, Stuart – thanks for everthing you've done to get me where I am today.
- And finally, this thesis is dedicated to my late mother, Denise.

Contents

Abstract **i**

Declaration **ii**

Acknowledgements **iii**

Contents **iv**

Index of Tables **vi**

Chapter One **1**

Introduction **1**

1.1 Overview 1

1.2 Synopsis of the Current Study 5

Chapter Two **6**

Cross-Cultural Ability Testing: The existing literature. **6**

2.1 Theoretical Approaches to Intelligence 6

2.1.1 Psychometrically-based Theories 6
2.1.2 Piagetian Theories 8
2.1.3 Information-processing Theories 9
2.1.4 The Theory of Multiple Intelligences 11

2.2 The Assessment of Ability in Multicultural Settings 12

2.2.1 Introduction 12
2.2.2 Cross-cultural Psychology 13
2.2.3 Cultural Bias in the Assessment of Ability 16
2.2.4 Aptitude vs. Achievement Tests 19
2.2.5 Methods of Establishing Cross-cultural Equivalence 20

2.3. The Raven's Coloured Progressive Matrices Test 24

2.3.1 Introduction 24
2.3.2 Validity 26
2.3.3 Reliability 28
2.3.4 The Application of the Raven's in African society 29
2.3.5 The Importance of this Study 33

Chapter Three **38**

Methodology **38**

3.1 Participants 38

3.2. The Instrument 39
3.2.1 Visual Aids 39
3.2.2 The Translated Instructions 40

3.3 Procedure 41

3.3.1 Data Collection 41
3.3.2 Data Processing and Analysis 42

Chapter Four 44

Results 44

4.1 Description of the Sample 45

4.2 Breakdown in terms of Age and Gender 45

4.2.1 The Mean Ages of the Participants. 47

4.3 The Scores on the Raven's CPM: Normative Sample 49

4.3.1 The Distribution of the Scores 49
4.3.2. The Mean, Standard Deviation and Range of Scores 51

4.4 Comparison of Scores across the Independent Variables 51

4.4.1 Comparison of Scores Across Preschool & Failed Grades 52
4.4.2 Comparison of Scores across Grades 52
4.4.3 Comparison of Scores across Gender 55
4.4.4 Comparison of Scores across Age 56

4.5 The Norms 56

Chapter Five 58

Discussion and Conclusion 59

5.1 Discussion of Results 59

5.1.1 The Effects of Age on Test Scores 60
5.1.2 The Effect of Gender on Test Scores. 61
5.1.3 The Effect of Grade on Test Scores 62
5.1.4 The Effect of Failed Grades and Preschool Attendance on Test Scores 63

5.2 The Norms 63

5.2.1 Comparison of Norms with Raven et al. (1990). 64

5.3 Limitations of the Study and Recommendations for Further Research 67

5.4 Conclusion 69

References 70

Appendix 1 76

Appendix 2 77

Appendix 3 78

Appendix 4 79

Index of Tables

Table 1: Sample - by GRADE **46**

Table 2: Sample - by AGE GROUPS **48**

Table 3: t-Test results – Comparison of differences in Ages between the Genders **49**

Table 4: Means, Standard Deviations, Range and Minimum and Maximum Scores by GRADE **53**

Table 5: t-test Results for the Comparison of scores across GENDER **55**

Table 6: Mean Scores, Standard Deviation, Range, Minimum & Maximum Scores by AGE **57**

Chapter One

Introduction

1.1 Overview

The 1996 Bill of Rights declaration 'Everyone has the right to a basic education' has led to an influx of previously 'disadvantaged' children into classrooms across South Africa, resulting in pupil-teacher ratios as high as 49:1 (Shindler, 1997). Recent revisions in state Education policy have led to the introduction of a new curriculum programme – Curriculum 2005, specially designed to meet the changing needs of our multicultural society. In addition to their contributions in the planning and evaluation of the new programme (Foxcroft, 1996), clinical psychologists are also depended upon for the fair and accurate cross-cultural psychometric assessment and placement of children within it. The ensuing increase in referrals from teachers and parents to clinical psychologists working in community outpatient settings throughout South Africa, must therefore be seen to be exerting considerable pressure on already over-taxed resources.

In a comprehensive review of contemporary concerns in mainstream South African cross-cultural assessment, Retief (1988) concluded that despite its historically racist origins, considerable progress has been made in eradicating regional (and internationally-recognised) obstacles to the practice of fair cross-cultural psychometric assessment. A review of the international literature revealed the most important obstacles identified as playing a role in any form of cross-cultural assessment to be: (1) language and translation (Bracken and Barona, 1991; Reddy, Knowles & Reddy, 1995); (2) a lack of familiarity with the testing situation and

process (Lonner, 1990; Merenda, 1994); (3) content, method and item cultural bias (Poortinga, 1995; Van de Vijver & Poortinga, 1995); and finally (4) a shortage of appropriate normative data (Geisinger, 1994; Helms-Lorenz & Van de Vijver, 1995).

In a 'blueprint for relevant assessment practices in South Africa', Foxcroft (1996) asserts:

> *"[t]hat a larger storehouse of appropriate measures for use in South Africa should be created by… developing culture-reduced tests with appropriate norms for diverse groups of children and adolescents in our country; and adapting and norming useful tests that have been developed overseas" (p. 12 –13).*

She argues further that the range of the term assessment should be broadened to include appropriately normed screening measures that could be used effectively in the testing of large groups of children. It is asserted that this approach will facilitate the more effective channeling of pupils through the new school curriculum, through:

> *"[t]he early identification of those children requiring an in-depth diagnostic assessment to verify developmental delays and plan for remedial intervention, …those requiring further monitoring and possible rescreening in the future, and finally, those requiring no further evaluation" (p. 7).*

Although developed and normed on British and American populations, the Raven's Coloured Progressive Matrices test (CPM) is internationally recognised as a culture fair test of non-verbal intelligence for young children (Raven, Court and Raven, 1990). It is important to note however, that the test cannot be employed to assess general intelligence, unless used in conjunction with an equally reliable and valid test of verbal intelligence (Raven et al., 1990). This easily administered, pencil and paper

test has no time limit and comprises three sets of twelve problems, arranged to "assess mental development up to a stage when a person is sufficiently able to reason by analogy to adopt this way of thinking as a consistent method of inference" (Raven et al., 1990, p. CPM2).

The test's innovative design thus effectively minimises and/or eliminates the four major cultural and educational obstacles to the construction of culture-reduced tests identified by Anastasi and Urbania (1997); namely language, test content, reading and speed. The Raven's CPM has been used effectively in a wide variety of cross-cultural settings for children between the ages of 5 and 11; and normative data has been successfully established for among others – Inuit children in Arctic Quebec (Wright, Taylor and Ruggiero, 1996); children in rural India (Barnabas, Kapur & Rao, 1995); and Polish children (Jaworska & Szustrowa, 1993). Although used extensively across a wide variety of settings in South Africa, there are currently no available South African norms for this test. This pilot study therefore seeks to make a positive contribution to the effective channeling of Xhosa-speaking primary school children in the peri-urban 'township' settings in the Grahamstown region, by developing much needed local normative data.

In a similar endeavor, Vass (1992) established normative data for the Raven's Standard Progressive Matrices Test (the version for individuals older than 11 years and 6 months) for a representative sample of Xhosa-speaking Secondary School pupils in the Grahamstown region. In addition, he also investigated the effect of test presentation on test results by employing two different methods of test presentation for his normative and alternate samples. Respondents in the smaller (alternate)

sample were administered the test using the original instructions translated into Xhosa, and two sets of visual aids (wherein the Record form and Item 1 of the test were reproduced double the size of the original as stipulated by Raven et al., 1990). In contrast, those in the normative sample were given the same set of instructions except that they had been extended to cover Item 2; and were presented with three sets of visual aids three times the size of the originals, where the six possible answers for each of the two test items were represented in the form of movable parts. The latter modification of method of test presentation allowed respondents to participate more actively during the 'instructional phase' of the test's administration (Vass, 1992).

His results showed that (1) the norms generated for the normative sample were considerably lower than previously established norms in similar studies conducted overseas; (2) that pupils in the normative sample scored significantly higher than those in the alternate sample (indicating that method of presentation on the Raven's SPM is important when assessing pupils that may be regarded as disadvantaged); (3) that male respondents scored significantly higher than female respondents; and finally, (4) that there was a significant effect of age, educational level and gender on test scores.

Thus to conclude, the primary objective of this study is to establish norms for the Raven's CPM test for Xhosa-speaking Primary School pupils in the Grahamstown region. Secondly, the results obtained will be compared with normative data established for this test in Western Populations; and examined across the different age

groups, genders and grades in order to identify the possible effects of these variables on test scores.

1.2 Synopsis of the Current Study

Chapter Two reviews the literature in the following areas: firstly, the various theoretical traditions that have been developed and adopted in the study of human intelligence; secondly, the practice of ability testing in multicultural societies, and the problems associated with this process; and finally, the administration of the Raven's CPM in multi-cultural settings including the limited South African literature.

Chapter Three begins with an overview of the methodology adopted in this study. In addition, the sampling procedures, method of test administration and statistical procedures used in the analysis of the data are outlined in detail.

Chapter Four comprises the presentation and analysis of the results obtained in this study, and is supplemented with tables and graphs wherein the findings are presented in a simpler and more easily accessible format.

Chapter Five summarises the main findings of the study and the implications of the results are discussed within the context of the both the literature review and the target population under study. Thereafter the limitations of the study are discussed and recommendations for further research are proposed.

Chapter Two

Cross-Cultural Ability Testing: The existing literature.

2.1 Theoretical Approaches to Intelligence

This review is not intended to provide a definitive answer to the question: "What is intelligence?" but rather to locate the Raven's CPM test within the context of the various approaches or traditions that have developed in the study of human intelligence. The theoretical foundations that underpin the construction and development of the Raven's Progressive Matrices series of tests were based in the work of Spearman (a pioneer in the psychometric tradition); who developed what he termed the 'two-factor theory of intelligence' (Raven, Court & Raven, 1993). Thus, for the purposes of this study, it is important to briefly explore this theory within the context of the various alternatives that have emerged in this area of psychological research.

2.1.1 Psychometrically-based Theories

Rust and Golombok (1989) assert that by 1910, the principles of psychometric theory had been established by Galton, Spearman and Pearson, and were being applied "almost entirely in the development of what had come to be called intelligence tests" (p. 5). The first intelligence tests were designed by Galton, who investigated individual differences in psycho-physical abilities (such as visual and auditory discrimination); reflecting his view that sensory discrimination and 'mental speed' were positively related to intelligence (Brody, 1992). Galton's approach to

intelligence theory has been described as 'bottom-up' in that his focus was on basic-level, mental processes (Sternberg, 1987).

Binet designed the first intelligence test to be used for the purposes of educational selection in 1905 (Rust & Golombok, 1989). Binet's theoretical approach asserted that complex mental functions such as comprehension, memory and imagination provided a more accurate indication of intellectual ability (Brody, 1992). Thus, in contrast to Galton, Binet adopted a 'top-down' theoretical approach in that he chose to focus on more complex, judgment-based higher-level mental processes (Sternberg, 1987). His test was later refined and its derivatives (such as the Stanford-Binet in America) gained popularity throughout the world in the diagnosis of mental handicap (Rust & Golombok, 1989).

During the same period, Spearman used the psychometric principles of factor analysis to develop a more comprehensive theory of intelligence. His 'two-factor' theory of intelligence assumed that all possible measures of intelligence could be divided into two separate components. The *general* or *'g'* component was defined by Spearman as being determined by "that which the measure has in common with all other measures of the common intellective function"; and the *specific* or *'s'* component, as being "that which is specific to each measure" (Brody, p. 3).

In addition, Spearman argued that the *'g'* factor itself was further comprised of two distinct components, namely *eductive* and *reproductive* ability. These were defined as follows:

> *"Eductive ability refers to the mental activity involved in making meaning 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 involving many mutually dependent variables". Reproductive ability refers to behavior involving the "mastering, recalling, and reproducing the (largely verbal) material which forms a cultural store of explicit, verbalised knowledge" (Raven et al., 1993, p. G1*).

In 1938, Raven utilised Spearman's theory as a rationale for test construction in the development of his Progressive Matrices test, which was designed in order to assess as accurately as possible the eductive ability component of *'g'* in individuals (Raven et al., 1993). The test itself is discussed in more detail at a later stage in this chapter.

Thus, in sum, the theoretical focus of the psychometric tradition was on accounting for the differences observed in individuals' intellectual abilities, and in developing theories that attempted to account for these 'psychometric' differences (Kail & Pellegrino, 1985).

2.1.2 Piagetian Theories

Piaget's approach to the study of intelligence reflected his view that any theory of intelligence would have to take into account both the stages in cognitive development, and the mechanisms by which this development took place (Brody, 1992). Within this tradition, Piagetian or Cognitive Developmental theorists focus on the similarities in intellectual growth found among all individuals (Kail & Pellegrino, 1985). Intellectual growth is here defined in terms of a set of stages that are "qualitatively different, reflecting major changes in cognitive structure" (Kail and Pellegrino, 1985, p. 157).

Piaget was thus opposed to the afore-mentioned "static normative view, where intelligence changes little over development" (Kail and Pellegrino, 1985, p.158), and highlighted the fact that the study of intelligence should be broadened to include the means by which individuals achieve intellectual growth. The dynamic role of the environment on the development of intellect was thus acknowledged and incorporated within this theoretical perspective; with the processes of maturation and experience viewed as playing a central role (Brody, 1992). Although the focus here is on commonalities or similarities, cognitive developmental theory also seeks to address the matter of "how individuals differ in terms of their intellectual growth" (Kail & Pellegrino, 1985, p. 159). It is important to note that Raven et al. (1990) refer to the work of Wiedl and Carlson (1976) wherein the CPM has been related to Piagetian conservation concepts, "noting a development in the reasoning processes required for CPM solutions from perceptual to conceptual" (p. CPM 22).

To conclude, this tradition is not primarily concerned with the data produced by intelligence tests, but rather opts to study intelligence along a developmental continuum. Most of the data supporting the cognitive-developmental theory is derived from the results of experiments, wherein individuals at varying points in their intellectual development are observed performing a wide variety of cognitive or intellectual tasks.

2.1.3 Information-processing Theories

The advent of computers in the early 1960's impacted significantly on the discipline of psychology, and human intelligence began to be studied using computer-based models of information processing. Within this theoretical tradition, the emphasis is

on both theoretical and empirical analyses of performance on both simple and complex tasks (Kail & Pellegrino, 1985). It also had a significant impact on the popular press, where the concept of 'artificial intelligence' was to receive considerable coverage[1].

More recently, Sternberg (1987) developed another information processing-based theory. His componential model of intelligence is known as the Triarchic theory of human intelligence. Its principle construct is the 'component', which is defined as:

> *"an elementary process that operates on internal representations of objects or symbols…thus, when a person sees an object, the sensory input is translated by such a component into a conceptual representation in the mind." (Rust and Golombok, 1989, p. 17).*

The Triarchic theory states that there are three kinds of information-processing components arranged to form a hierarchy, namely; (1) metacomponents or 'higher order' executive processes; (2) performance components or 'lower order' components which execute the instructions of the components; and (3) knowledge acquisition components such as selective encoding, selective combination and selective comparison (Sternberg, 1987).

Another more recent example of this approach is the Pass (planning, attention, simultaneous and successive) theory of intelligence developed by Kirby and Das (1990). The Pass theory states that cognition can be conceived as the outcome of three interconnected neurological systems, namely centers for arousal and attention, processing or encoding processes and planning (Kirby & Das, 1990). Of relevance

[1] The term refers to attempts by psychologists to design a computer that would replicate the functioning of the human brain, and thus possess a 'mind of it's own'.

here, is that Raven et al. (1990) note the separation within the theory of the processing function into simultaneous and successive processing and argue that the Raven's CPM "has proved repeatedly to load highly on the simultaneous processing factor, so that it has become the criterion instrument for this purpose (typically with loadings from .75 - .85)" (p. CPM 22).

Thus to conclude, this tradition generally defines intelligence in terms of dynamic and multifactorial thought processes or relationships which, by definition, are occurring simultaneously during any one type of intellectual activity[2]. The information-processing theoretical tradition is therefore acknowledged as making a useful contribution to the study of intelligence, in that it represents an innovative attempt at integrating earlier, conflicting theoretical traditions (Kail & Pellegrino, 1985).

2.1.4 The Theory of Multiple Intelligences

Finally, almost fifty years after the Raven's Progressive Matrices test was first developed, new theories regarding the nature of human intelligence continue to surface. Most recently, Gardner (1983) put forward what he termed the theory of multiple 'intelligences'. His theory asserts that there are in fact, six distinctly different types of intelligence that have been identified. Gardner drew upon a wide array of empirical evidence to support the existence of the following distinct types of intelligence: musical; bodily-kinesthetic; personal or 'intra-personal'; linguistic; logical-mathematical intelligence; and finally, spatial (Kail and Pellegrino, 1985).

[2] For a theoretical account of information-processing involved in the Raven's Progressive Matrices tests, see Carpenter, Just and Shell (1990.)

Despite receiving a somewhat mixed reception, Gardner's theory has however also been acknowledged as yet another resourceful attempt to integrate pre-existing theoretical traditions, and to answer the question that continues to plague psychologists today, namely: 'what is intelligence?' (Sternberg and Detterman, 1986, cited in Jensen, 1998).

2.2 The Assessment of Ability in Multicultural Settings

2.2.1 Introduction

In this section, the literature regarding the psychometric assessment of abilities is reviewed. The target population of the current study is acknowledged as being deeply embedded within a society characterised by a kaleidoscope of cultural groups, 11 official national languages and a history of deeply entrenched political oppression and racial discrimination. For this reason, this section of the literature review will be firmly located within the context of the difficulties associated with culturally relevant test usage in multicultural societies such as ours. The history of the discipline of cross-cultural psychology will be briefly examined to allow for the drawing of a distinction between its two main traditions namely, *cross-cultural comparison* and *culture-fair test use*. Thereafter, the term 'cultural bias' is defined in the context of ability assessment; and the various forms of cultural bias are outlined. A second important distinction will be drawn between the two different types of ability tests available, namely: aptitude and achievement tests. Finally, this is followed by an exploration of the diverse approaches that have been adopted to eliminate or overcome the challenges inherent in the fair assessment of ability in multicultural societies.

2.2.2 Cross-cultural Psychology

Yau-Fai Ho (1994) offers the following definition of Cross-cultural Psychology:

> *"[t]he scientific study of human behaviour and mental processes, including both their variability and invariance, under diverse cultural conditions; with it's primary aims being to investigate the systematic relations between behavioural variables and ethnic-cultural variables, and the generalisations of psychological principles" (p. 4).*

Although the discipline was only formally recognised some thirty years ago, a review of the literature reveals that many of its formative influences can be traced back much further than that.

In a paper entitled 'Cultural Bias in Assessment: Historical and Thematic Issues', Poortinga (1995) provides an excellent overview of historical developments within this discipline. According to Poortinga, Cross-cultural psychology can be viewed as comprising two distinct traditions, with one having developed later than the other[3]. The first tradition originated in early cross-cultural comparative studies, wherein the primary concern was theoretical questions about cultural differences in human behaviour. The second developed in response to the first and is referred to in the literature as the tradition of culture-fair testing.

He argues further that the *cross-cultural tradition* developed at the turn of the century, when researchers began using newly developed psychometric assessment procedures to investigate 'theoretically driven' research hypotheses (concerning human behaviour

[3] Unless otherwise stated, this section of the discussion is based primarily on this text.

and cognition) across different cultural populations and geographical regions. These studies were generally based on the following assumption:

> *"[t]hat the operationalisation of a trait in an assessment procedure is possible in such a way that the performance of the testee is determined by individual talent, independent of cultural context" (p. 140).*

Furthermore, although a broad range of differences in behaviour and cognition were under investigation during this time period, it was the area of intelligence that was to eventually prove most popular.

The popularity of the assessment of intelligence among cross-cultural comparative researchers has been attributed to a number of interesting historical factors. Initially, intelligence tests themselves increased in popularity when Binet produced the first test for the purposes of educational selection, in 1908 (Jensen, 1980). Intelligence tests were subsequently used extensively by the military (approximately 10 million individuals were tested during World War Two), and in educational settings such as schools and universities (Maloney & Ward, 1976). Both laymen and academics from a wide range of disciplines were thus familiar with the concepts of IQ and intelligence testing. In response, cross-cultural psychologists intensified their efforts to develop and refine psychometric assessment instruments for the testing of theoretical hypotheses pertaining to question of the 'universality' of human intelligence.

Later, this focus was further intensified by the increasing criticism generated by intelligence tests (or 'IQ tests' as they had become known) as a result of their rapid and widespread application throughout all levels of modern society. By the early 1950's, articles attacking intelligence tests had begun appearing regularly in both the academic and popular press, as the social and political implications of their use

became increasingly apparent. Academics and laymen alike now began to object not only to the tests themselves, but also to the practice of intelligence testing in general (Maloney & Ward, 1976). In the United States of America, intelligence tests made their way into the Courts, where the issue of discrimination (on the basis of 'biased' IQ test results) became the subject of numerous lawsuits (Jensen, 1980). It is interesting to note that the majority of these cases involved the allocation of Black and Hispanic children to special education classes on the basis of IQ test results (Maloney & Ward, 1976).

In addition, the new methodological and psychometric procedures that had been developed were now being used to facilitate the more accurate identification of the psychological effects of numerous sociological or socio-cultural factors on test performance (Cronbach & Drenth, 1972 cited in Poortinga, 1995). Attempts to construct a culture-free test were aborted in the face of the overwhelming evidence to the contrary; and the problem of test bias (gender, racial or cultural group) became the central theme in the studies that followed. The umbrella term 'cultural bias' was introduced to encompass the many, ever-increasing dimensions along which the tests were being discovered to be culturally biased (Jensen, 1980). The cumulative effects of these theoretical and methodological developments led to a greater awareness of the pervasive influence and diverse effects of culture on test performance and to its recognition as the single most pressing methodological problem facing cross-cultural studies. Thus, according to Poortinga, by the early 1960's a new *tradition of culture-fair testing* had emerged wherein efforts were concentrated on developing strategies to eliminate or overcome obstacles to the fair assessment of ability.

Thus to conclude, while the primary aim of the new 'culture-fair tradition' was the "fair and ethnically non-discriminatory use of tests within a single society"; the cross-cultural comparative tradition was concerned instead with "whether the scale formed by the score variable reflects the trait of interest in the same way in each of the cultural populations" (Helms-Lorenz & Van de Vijver; 1995, p. 192).

2.2.3 Cultural Bias in the Assessment of Ability

Cultural bias can thus be broadly defined as the "collective effect of all unintended variables in cross-cultural studies" (Poortinga, 1995, p. 141). The term is used to refer to problems posed by the presence of socio-cultural differences between two sample populations. A review of the international literature reveals cultural bias as being evidenced in the following four general areas: (1) language and translation (Reddy, Knowles & Reddy, 1995); (2) a lack of familiarity with the testing situation and process; (3) content, method and item cultural bias (Poortinga, 1995; Van de Vijver & Poortinga, 1995); and finally, (4) a shortage of appropriate normative data (Geisinger, 1994; Helms-Lorenz & Van de Vijver, 1995). In their comprehensive evaluation of the role of cognitive tests in multicultural school settings, Helms-Lorenz and Van de Vijver (1995) present an overview of the three main types of cultural bias that have been identified namely: construct, method and item bias. In addition, two potential sources of these types of bias were identified, namely 'subject-related' factors and what is termed the 'cultural loading' of the test.

Construct bias refers to the question "How does one know whether a measure adapted to a new language and/or culture measures the same construct or constructs that it did in the first?" (Geisinger, 1994, p. 304). Poortinga (1995) asserts that this more

abstract form of bias stems from the assumption that "existing (western) conceptualisations of psychological constructs or traits in one cultural population are relevant also for other quite different cultural (or racial) populations" (p. 140). In contrast, the term 'method bias' generally refers to anomalies at the test level, and is defined as "the influence of a cultural factor on the test scores, such as differential stimulus familiarity that is shared by most or even all items" (Helms-Lorenz & Van de Vijver, 1995, p.159). Method of test presentation may also represent a source of method bias (Vass, 1992). Other examples of this form of cultural bias include a lack of familiarity with timed tests and the multiple-choice testing format. Finally, item bias refers to anomalies at the item level that are "systematic though unintentional, such as poor item translation" (Helms-Lorenz & Van de Vijver, 1995, p.159) and is believed to be present whenever "persons from different cultural groups with the same ability level have an unequal probability of responding correctly to an item" (Helms-Lorenz & Van de Vijver, 1995, p.164).

Helms-Lorenz and Van de Vijver's concept of cultural loading is defined as referring to "the explicit or implicit reference to a specific cultural context - usually the culture of the test composer, in the instrument or its administration"; and thus encompasses all three forms of bias (p. 160). Furthermore, the authors have identified five potential sources of cultural loading. These include:

> *"[t]he tester (when the tester and testee are of a different cultural background; the testees (i.e. intergroup differences in educational background, scholastic knowledge and test-wiseness); the tester–testee interaction (e.g. communication problems); response procedures (e.g. differential familiarity with time limits in test procedures); and finally, cultural loadings in*

> *the stimuli (such as differential suitability of items for different cultural groups due to stimulus familiarity)" (p. 160).*

It is further asserted that cultural loadings are not an intrinsic property of the instrument, but will depend on the characteristics of both the instrument and the target sample.

Problematic characteristics or traits within the testees themselves represent another source of potential bias, and are collectively referred to as 'subject-related factors' (Helms-Lorenz & Van de Vijver, 1995). These are listed as:

> *"[t]he testee's verbal ability (which may be compromised by item phrasings in the test or it's instructions which included culturally biased idioms or words); the dominant cultural norms and values in the testee's cultural group; the fact that testees may differ in their level of test-wiseness (such as a lack of familiarity with multiple choice tests or time limits); and finally the target population may differ in terms of their acculturation strategy and their expectations for their future" (p. 161-162).*

To conclude, it is therefore theoretically possible for the above-mentioned factors to be a source of all three forms of bias, although their probabilities are significantly different. Helms-Lorenz and Van de Vijver illustrate this point by citing the example of construct bias, which "is far less likely to be a problem in school achievement tests than in aptitude tests" (p. 162). Finally, the significance of these factors (and the various forms of cultural bias) in the context of the Raven's CPM test itself is outlined in the following section.

2.2.4 Aptitude vs. Achievement Tests

This review also revealed the single most important factor in the determination of test bias (aside from the presence of construct, method or item bias) to be the intended goal of the testing procedure itself; or stated even more simply, the inferences or generalisations that are to be made on the basis of test results (Helms-Lorenz & Van de Vijver, 1995; Poortinga, 1995). Ability tests are "regularly applied in primary schools in many countries for, among other things, progress testing, assessment of learning difficulties, school advice and vocational guidance" (Helms-Lorenz & Van de Vijver, 1995, p.160). It is therefore important at this juncture, to note the two different types of ability tests currently available, as well as the differences between them (Helms-Lorenz & Van De Vijver, 1995).

School achievement tests are principally designed to measure the intellectual knowledge and skills acquired through education (Helms-Lorenz & Van de Vijver, 1995). This type of intelligence has been referred to as *crystallised* intelligence (Cattell & Butcher, 1968 cited in Helms-Lorenz & Van de Vijver, 1995). In contrast, aptitude tests are purported to be the least reliant on previous learning experiences, and to assess what is referred to by Cattell and Butcher (1968) as *fluid* intelligence (cited in Helms-Lorenz & Van de Vijver, 1995). Learning potential tests and intelligence tests are both purported to be good examples of aptitude tests (Helms-Lorenz & Van de Vijver, 1995).

Some researchers are however of the opinion that most traditional intelligence tests usually involve some verbal ability and scholastic knowledge, and therefore measure a mixture of crystallised and fluid intelligence (Cattell & Butcher, 1968, cited in

Helms-Lorenz & Van de Vijver, 1995). Others are more positive, pointing out that there are some intelligence tests which are "fairly close to aptitude tests", and the Raven's tests are cited as an example wherein "a deliberate attempt was made to use simple stimulus material that is not acquired in school" (Helms-Lorenz & Van de Vijver, 1995, p.160).

2.2.5 Methods of Establishing Cross-cultural Equivalence

There are those such as Frijda and Jahoda (1966), who choose to adopt a culturally relativist approach towards the problem of bias in tests, arguing that "stimulus material will always be susceptible to differences in cultural backgrounds of testees" (cited in Helms-Lorenz & Van de Vijver, 1995, p.161). However, a review of the international literature shows that the counteractive measures developed to minimise and/or eliminate test bias can be roughly classified into five main approaches or strategies. These include: (1) the adaption of existing tests; (2) the construction of new, more culturally appropriate assessment instruments; (3) the application of differential norms; (4) the design of culture comparative studies that make less strict demands on equivalence by avoiding direct comparison of scores on a single test; and (5) the psychometric control of bias (Helms-Lorenz & Van de Vijver, 1995; Poortinga, 1995). Each of these strategies will now be briefly discussed in turn.

The first approach to eliminating cultural bias involves the adaption of existing tests, and is by far the most popular. Proponents of this approach submit existing standard tests to "analyses of the construct underlying the test, an evaluation of the adequacy and application of the construct in the target group, and the adequacy of operationalisations in the items of the instrument" (Hambleton, 1994, cited in Helms-

Lorenz and Van de Vijver, 1995, p.163); and utilize specific, well-researched procedures, such as the back translation method (Brislin, 1980), for translation of the test into a second language.

Geisinger (1994) views the shift in terminology in the literature (from translation to adaption) as reflecting the acknowledgment of the multiple dimensions along which a test may require modification; and the determination of this process by *qualitative* differences between the two populations. Typically, assessment devices are adapted for new settings because a need for that measure is apparent in the target culture (Geisinger, 1994). Thus, the adaption procedures for any assessment measure will always depend on the idiosyncratic characteristics of the target population or cultural group. He points out that in some cases translation into another language may be more important, while others may require the adaption of cultural content, and then there are those cases that will require both.

Several guidelines for the translation and adaption of existing tests (wherein the required procedures are clearly outlined for each step in the process) are also available (see Bracken & Barona, 1991 and Geisinger, 1995). In addition, Van de Vijver and Leung (1995) recommend, "providing clear and lengthy instructions, exercises after the examples and avoiding complicated grammatical structures and local idioms" (cited in Helms-Lorenz & Van de Vijver, 1995, p.163). In their review of the literature, Helms-Lorenz and Van de Vijver (1995) note that most test translations reported are "aimed at enhancing the validity of an instrument for a particular group; and are therefore much less complicated than the few aimed at enhancing a test's validity in a multicultural setting" (p. 163). It is important to note however, that

whenever any item content in a test is altered, the comparison of tests scores across the two different populations is no longer possible (Poortinga, 1995).

The second approach to the problem of cultural bias involves the construction of new, more culturally appropriate tests altogether (Helms-Lorenz & Van de Vijver, 1995; Poortinga, 1995). Cattell's Culture-fair Intelligence tests (Cattell & Cattell, 1963) and the Raven's Progressive Matrices are cited as good examples of these types of tests, as they were products of the culture-free and culture-fair test movements respectively (cited in Helms-Lorenz and Van de Vijver, 1995). However, this is generally a very expensive and labour intensive alternative, requiring a lengthy and protracted process of test construction, validation and standardisation and is therefore a less popular alternative.

The third method involves the use of different norms for existing tests and has been labeled a "politically sensitive" option (Helms-Lorenz & Van de Vijver, 1994, p.162). Although the original version of the test is used and scored in the same way, the scores receive different interpretations due to the application of differential (depending on the testee's cultural group membership) sets of norms (Helms-Lorenz and Van de Vijver, 1994). This approach is usually used to compensate for social inequalities such as unequal opportunities in job application processes; with examples ranging from "choosing different pass-fail cutoff points for different cultural groups, to designating beforehand a fixed percentage of a cultural group to be appointed regardless of the average level of the group" (Helms-Lorenz and Van de Vijver, 1994, p.163). Social or political development programmes such as positive discrimination,

equal opportunities and affirmative action therefore favour this approach (Helms-Lorenz & Van de Vijver, 1994).

The fourth strategy used to counteract cultural bias involves designing culture comparative studies that make "less strict demands on equivalence…by avoiding direct comparison of scores on a single test…as the information required can be obtained from patterns of scores over measurement occasions" (Poortinga, 1995, p.141). This approach is utilised within the cross-cultural tradition and illustrated by a study designed by Irvine (1979) wherein he investigated the question of whether the factor analytic structure of intelligence is the same in various regions of the world (cited in Poortinga, 1995).

The fifth and final approach deals with cultural bias in the assessment of ability directly; "by identifying its effects statistically and correcting the instrument in such a way that it meets psychometric standards of equivalence" (Poortinga, 1995, p.141). An example of the psychometric control of bias is the statistical procedure known as item bias analysis or differential item functioning (Helms-Lorenz & Van de Vijver, 1994). Other examples include the analysis of translation equivalence, a procedure based on the back translation method developed by Brislin (1980). Poortinga (1995, p.145) points out that "most new developments seem to recognise that cultural bias is not some form of measurement error, but a systematic component of the variance that needs to be explained".

To conclude, a review of the international literature revealed the question of a test's suitability as depending on the degree to which it is free of content, method and item

cultural bias; and more importantly, on the intended purpose of the test scores (Helms-Lorenz & Van de Vijver, 1995; Poortinga, 1995). The three afore-mentioned types of bias can occur in all types of tests (though not with the same probability), with method bias being the most likely as:

> *"[m]ost subject related factors such as intergroup differences in verbal skills or test-wiseness will affect all items in a more or less uniform way, thereby inducing intergroup differences in average test performance that cannot be attributed to the construct of the test" (Poortinga, 1995, p.166).*

2.3. The Raven's Coloured Progressive Matrices Test

2.3.1 Introduction

The Raven's CPM is internationally recognised as a culture-fair or culture reduced test of non-verbal intelligence for young children (Raven et al., 1990). This easily administered, multiple-choice pencil and paper test has no time limit, and comprises three sets of twelve matrix designs arranged to "assess mental development up to a stage when a person is sufficiently able to reason by analogy to adopt this way of thinking as a consistent method of inference" (Raven et al., 1993, p. CPM2). In this version of the Raven's Progressive Matrices however, each item is printed with a brightly coloured background, making the test more appealing for children.

The testee is shown a series of patterns with parts missing. The parts removed are of simple shape and have been placed below the matrix, among other similarly shaped pieces (although the figures on those pieces do not compete the pattern) (Martin & Wiechers, 1954). The problems are easy to begin with, but grow more difficult as the

test proceeds "because the figures in the patterns to be completed remain simple but the relations between the figures become increasing complex" (Martin and Wiechers, 1954, p.143). The testee can either point to the pattern piece s/he has selected or write its corresponding number on the record form (Lezak, 1995). The total score is the total number of matrices completed correctly, and the test is thus scored out of 36.

As stated earlier, the Raven's Progressive Matrices (Standard, Coloured and Advanced) are the most widely used and extensively researched of the culture-reduced tests (Jensen, 1980). The Raven's CPM was initially standardised in the United Kingdom on a population of 627 Scottish school children in 1949 (but was later restandardised in 1982); and was recently standardised in the United States of America (Raven et al., 1990). The published norms for the Raven's CPM are presented in the form of percentiles, and are available for children aged between 5 ½ to 11 ½ years (Raven et al., 1990). Half-yearly age intervals were selected because mental development in childhood is viewed here as being "more like salmon leaps in the stream of life than the equally arranged lungs in a ladder", and this test was designed to assess such 'leaps' (Raven et al., 1990, p. CPM2). The child's mental age is thus the age at which the median score is equal to his or her raw score. The test is also available in booklet and board form (the latter being used with very young children), and can be administered to individuals or groups (Raven, et al., 1990).

A separate publication is available from the Publishers of the Matrices test wherein the literature regarding this test and the vast databank of normative data accumulated thus far is reviewed; and comprises over 2000 studies (Anastasi & Urbania, 1997). As a review of this literature is beyond the scope of this study, this review will be

limited to the test's validity and reliability and studies concerning it's application in African societies.

2.3.2 Validity

As discussed earlier, the rationale behind the test's construction was Spearman's two-factor theory of intelligence. The Raven's CPM thus purports to asses a unitary trait namely 'g', or more specifically, the eductive component of intelligence. [4] This type of intelligence has also been referred as analytic intelligence or the ability to "deal with novelty, to adapt one's thinking to a new cognitive perspective (Carpenter, Just and Shell, 1990). In terms of construct validity, there has been considerable controversy in the literature over the past few decades regarding 'g'; in particular, there has been some debate as to which factor(s) can be regarded as constituting 'g' (Carpenter et al., 1990) and whether it is in fact a unitary trait [5]. Nevertheless, Jensen (1998) argues that:

> *"… when the Progressive Matrices test is factor analysed along with a variety of other tests, it is typically among the two or three tests having the highest 'g' loadings, usually around .80 and … its most distinctive feature is its very low loadings on any factor other than 'g'." (p. 38).*

Martin and Wiechers (1954) sought to investigate the degree of correlation between the Raven's CPM and the WISC (Wechsler Intelligence Scale for children). Their results revealed correlations of .91, .84, and .83 between the score and the WISC Full Scale, Verbal and Performance IQ's respectively. In addition, the test correlated

[4] As discussed previously, eductive ability refers to "the mental activity involved in making meaning 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 involving many mutually dependent variables. (Raven et al., 1993, p.G1).
[5]See Jensen (1998) where this ongoing complex theoretical debate is covered extensively.

highest with Block Design (.74) and lowest with Information (.47), suggesting that the test is a culturally reduced measure of nonverbalised abstract thinking (Martin and Wiechers, 1954).

However, it important to note that cross-cultural research has shown that African children (and especially those from rural areas) usually perform poorly on Western pattern reproduction tests such as the Block Design assembly test, the Bender Gestalt test and the Draw-a-Person test (Serpell, 1994). In a cross-cultural study, Serpell sought to investigate this phenomenon, in the context of his observations of superior pattern reproduction skills in African children in the form of fully-functioning reproductions of "skeletal model cars built from scraps of wires" (1979, cited in Lonner & Malpass, 1994, p. 159). He compared the performance of children from two low-income neighbourhoods (in Zambia and Manchester, England) on a drawing pattern reproduction test, a wire-modelling pattern reproduction test and a clay-modelling pattern reproduction test. As predicted the English children performed better on the drawing task and the Zambian children out-performed the English children on the wire-modelling task, with no group difference on the clay modeling task . Serpell concluded that:

> "*the abstract psychological function of pattern reproduction can manifest itself in different ways according to the demands of the eco-cultural niche to which the subject's behaviour is adapted. In an environment abundant with paper and pencils, children would be likely to acquire greater pattern reproduction skills in the medium of drawing than in an environment where the relevant materials were scarce. Conversely, in an environment where modeling in wire was a recurrent play activity, children would be likely to acquire greater pattern reproduction skills in the medium of wire-modelling*

than in an environment where the activity was unknown (1979, cited in Lonner & Malpass, 1994, p. 160).

This study thus highlights not only the kind of activity required by the Raven's CPM test, but also the children in the target population under study's lack of exposure to it.

The test has generally demonstrated a low predictive validity when compared with school achievement tests, and Freeman (1960) argues that this may be due to the fact that the Raven's Progressive Matrices test "measures only a specific ability of a spatial perception factor, whereas achievement examinations require simultaneous use of several abilities among which verbal comprehension, verbal reasoning and numerical faculty may be important" (cited in Ogunlade, 1978, p. 467). Finally, and most importantly, this review of the literature has revealed that it is generally been agreed that when the test is used in populations other than those upon which the test has been standardised, local norms are considered best (Sigmon, 1983).

2.3.3 Reliability

In the recent standardisations, the retest reliability of the Raven's CPM was revealed to be .90 over the whole range of development (Raven et al., 1990). A split half-reliability estimate of .90 was obtained, with no differences found between ethnicity (Anglo, Black and Hispanic) or genders in a study conducted by Jensen in 1974 (cited in Raven et al., 1990). In a subsequent study by Carlson and Jensen (1981), the split half reliability estimate of .85 was established; with the estimates at ages 6, 7 and 8 generating estimates of .65, .86 and .85 respectively. There has however, been some evidence that when the test is administered to very young children, the CPM generates

lower reliability estimates (Sattler, 1982). Valencia (1984) explored the test's reliability for Anglo and Mexican-American schoolchildren in Grade Three and found it to be acceptably high and equal for both cultural groups. Valencia (1984, p.51) concludes that the practicing school psychologist consider the Raven's CPM as "an assessment tool when the intellectual abilities of children who are linguistically and culturally diverse in backgrounds are in question".

2.3.4 The Application of the Raven's in African society

Ogunlade's review of the African literature in 1978 concluded that the Raven's Progressive Matrices test (i.e. both the SPM and CPM) is highly suited to Third World populations. It must be noted however, that in the majority of African literature, the Raven's SPM (the version designed for children attending secondary school) is usually favoured. This version has been effectively used in Nigeria (Ogunlade, 1978); Tanzania (Klingelhofer, 1967) and Southern Rhodesia (Irvine, 1963) and West Africa (Berlioz, 1955) (cited in Ogunlade, 1978); Uganda (Silvey, 1963, cited in Klingelhofer, 1967); and South Africa (Notcutt, 1949 and Vass, 1992). As the majority of the studies reviewed employed the Raven's SPM and are also more than thirty years old, the decision was made to limit the review to the South African literature. It is argued that this literature is the most appropriate and relevant in the context of the aims of the current study.

Given the above-mentioned statements regarding the suitability and popularity of the Raven's Progressive Matrices tests in Third world countries; a review of the South African literature reveals a surprising paucity of research conducted on all three versions of the Raven's Progressive Matrices Test. To date, no local normative data

for the Raven's CPM has been published, mirroring the situation reported above in other African countries. Furthermore, although there have been some studies conducted using the Raven's SPM, the majority of this research was conducted in the 1950's and 60's.

Notcutt conducted the first South African study in 1949, wherein he attempted to standardise the Raven's SPM version of the test for Zulu-speaking school children (age between 11 and 15 years) living under urban and peri-urban conditions in Durban. He found that "over an age range where Raven's English sample shows a marked negative skewing, the Zulu sample showed an equally pronounced positive skewing" (Notcutt, 1949, p.68). However, he argued that if one chose to equate the median scores instead of the ages of the two samples, the distributions were really very similar. He found that the median score on the test for the Zulu sample at 11 ½ years was almost exactly the same as that of the English sample at 8 years; with Zulu 12 ½ equal to English 8 ½; Zulu 13 ½ equals English 9 years; and Zulu 14 ½ years only slightly lower than the English 9 ½ years. Notcutt (1949) also noted that when deviations from the median score were compared, each group's distribution of scores is similar to the other in that, there is a change from a positive skewing where the median score is 18 points, to a normal distribution where the median is about 25 points. From this comparison he concluded that the form of a distribution of an age group of the sample of these scores depends more on the test than on the group tested. Notcutt (1949) asserted further that:

> "*To increase one's score on the RPM from 44 to 50 is more difficult task than to increase it from 24 to 30, hence the distribution negative skewing at higher levels. Thus, the English norms are negatively*

> *skewed at ages where the Zulu norms (with a much lower average score) are positively skewed" (p. 70).*

Over the next four decades, the few South African studies involving the Raven's SPM focused solely on developing norms for adult populations. In 1951, Malan focused on developing Raven's SPM norms for a population of Afrikaans-speaking adults. Continuing the trend, Labuschagne (1955) established Raven's SPM normative data for adult natives of Central Nyasaland. In addition, both of the above sets of normative data have remained unpublished.

To date, only one South African researcher has attempted to develop Raven's SPM norms for school-age children. Vass (1992) established norms for a representative sample of Xhosa-speaking Secondary School pupils in the Grahamstown region. In addition, the effect of test presentation on scores was also investigated by employing two different methods of test presentation for a normative and alternative sample. Vass' method of test presentation for the normative sample combined the work of Pons (1974), who developed an alternative method of test presentation for the Raven's SPM, Crawford-Nutt (1976) and Schwarz (1963)(cited in Vass, 1992). Crawford-Nutt noted "that in the study of differences in intelligence test scores between blacks and whites the notion of test administration had largely been ignored" (cited in Vass, 1992, p.51) and argued that these reported differences in scores did not occur when the alternate method of test presentation for the Raven's was used; while Schwarz (1963) focused on using visual aids, supplemented by active demonstration (cited in Vass, 1992).

Respondents in Vass's smaller alternative sample were administered the test using the original instructions translated into Xhosa; and visual aids, wherein the record form and Item 1 of the test were reproduced double the size of the original (as stipulated by Raven et al., 1990). In contrast, the normative sample were given the same set of translated instructions, but were instead presented with visual aids three times the size of the originals (with the six possible answers for each of the two test items represented in the form of movable parts); with both the visual aids and the instructions extended to cover Item 2. These modifications allowed respondents to actively participate during the 'instructional phase' of test administration. Vass' results showed that (1) the norms generated for the normative sample were considerably lower than previously established norms in similar studies overseas; (2) that pupils in the normative sample scored significantly higher than those in the alternative sample (indicating that method of presentation on the Raven's SPM is important when assessing pupils that may be regarded as 'disadvantaged'); (3) that male respondents scored significantly higher than female respondents; and finally, (4) that there were significant effects of age, educational standard and education on test scores. Vass' primary contribution however, was the extension of Crawford-Nutt's finding to include "those respondents to the Raven's SPM that may be classed as disadvantaged" (1992, p.101).

There is only one South African reference to the Raven's CPM in the South African literature and the results of this study also remain unpublished. In 1994, Maree investigated the performance of mentally impaired adult Coloured patients on both the Raven's CPM and the South African Wechsler Adult Intelligence Scale (SAWAIS). Maree found that although for greater clinical accuracy the use of the SAWAIS is still

indicated, the Raven's CPM could replace the SAWAIS as a rough measure of general intellectual ability in adult clinical populations.

To conclude, despite the fact that the Raven's CPM is used extensively in educational and clinical settings throughout the country, there are currently no available locally generated norms for non-adult or clinical populations.

2.3.5 The Importance of this Study

In a comprehensive review of contemporary concerns in mainstream South African cross-cultural assessment, Retief (1988) concluded that despite its historically racist origins, considerable progress has been made in eradicating regional (and internationally-recognised) obstacles to the practice of fair cross-cultural psychometric assessment. In a more recent 'blueprint for relevant assessment practices in South Africa', Foxcroft emphasized the urgent need for the development of (1) "culture-reduced tests with appropriate norms for diverse groups of children and adolescents in our country" and for (2) "the adaption and norming of useful tests that have been developed overseas" (1996, p.12-13). In addition, Foxcroft advocates the extension of the range of the term assessment to include appropriately normed screening measures that could be used effectively in the testing of large groups of children and for the monitoring of their progress with the new program. It is asserted that these will facilitate the more effective channeling of pupils through the new school curriculum, through:

> *"[t]he early identification of those children requiring an in-depth diagnostic assessment to verify developmental delays and plan for remedial intervention, …those requiring further monitoring and possible*

rescreening in the future, and finally, those requiring no further evaluation" (p. 7).

Writing from a clinical perspective, Shuttleworth-Jordan (1996) writes against the overly hasty abandonment of existing tests in South Africa and the "attitude of nihilism…which occurs because tests have not been designed for application among a particular population, or because appropriate normative data is not yet available" (p. 96). She argues further that although this stance has relevance when dealing with rural and illiterate or semi-literate populations, the processes of acculturation taking place within urbanised South African populations must be acknowledged as having a strong mediating influence on test performance. In addition, Shuttleworth-Jordan emphasises the importance of distinguishing between the following:

> *"(1) Racial differences (i.e. ethnic factors), which in themselves may be considered to cause quantitative changes in cognitive test performance, although this remains an unresolved and highly contentious issue (see Lezak, 1995); and*
>
> *(2) Socio-cultural differences (i.e. factors such as preschool socialisation experiences, primary language, current language usage, levels of education, socio-economic status, and test sophistication) which are frequently associated with racial differences, and which are known to account for significant variations in test performance (Ardila, 1995; Lezak, 1995; cited in Shuttleworth-Jordan, 1996, p.96)".*

She also cites the research of Avis (1995) and Dyall (1996), where an absence of differential effects on standard psychological tests between samples of black and white South African children that had been matched for age and level of education was found. Significant improvements in test performance on urban relative to rural black groups, particularly among more educationally advantaged black populations

who are studying in the medium of English have also been found (Viljoen, Levett, Tredoux & Anderson, 1994, cited in Shuttleworth-Jordan, 1996).

According to Shuttleworth-Jordan, not only does the existing outmoded 'normative-based abilities test paradigm' (wherein people are categorised purely on the basis of the test results) fail to take into account the dynamic nature of socio-cultural influences, but it also ignores certain internationally recognised commonalities in brain-behaviour relationships. Thus, Shuttleworth-Jordan agrees that a generally broader-based model of ability assessment is needed in South Africa, suggesting that:

> *"[t]est results form only part of the overall data base, alongside clinical history, the clinical presentation, and the pattern and processes of cognitive performance across a series of functional domains, in the search for a conceptually coherent set of diagnostic criteria" (p. 96).*

Recent trends in contemporary South African psychological assessment can therefore be described as reflecting the many difficulties inherent in the practice of ability assessment in multi-cultural settings. The traditional narrow 'normative-based' approach is viewed as being hopelessly ineffective in dealing with the multifarious demands of our culturally pluralistic society, and a general broadening of perspective, scope and approach to ability assessment is being called for. It is therefore argued here that the Raven's CPM represents an especially effective assessment tool in this context. Furthermore, it is asserted that the test is capable of meeting the above-mentioned pressing and idiosyncratic assessment demands for the following reasons:

1. As noted above, the Raven's Progressive Matrices tests are highly suited to African populations (Ogunlade, 1978); and the Raven's CPM has been

internationally recognised as a reliable culture-reduced test of non-verbal intelligence for young children (Raven et al., 1990).

2. The review of the literature has shown that the Raven's CPM not only has a sound theoretical basis underlying its construction and development, but also offers an extensive and well-established research database which has accumulated throughout the world over the past 50 years.
3. In addition, the test's developers have not only ensured that the test makes significant contributions towards eradicating the afore-mentioned threats posed by the problem of cultural bias, but its innovative design also effectively minimises and/or eliminates the four major cultural and educational obstacles to the construction of culture-reduced tests identified by Anastasi and Urbania (1997); namely language, test content, reading and speed.
4. The review of the relevant literature has also shown the Raven's CPM to embody a method of establishing cross-cultural equivalence in the assessment of intelligence that is well suited to the population currently under study.
5. Furthermore, the Raven's CPM can be administered to either groups or individuals and although there is no time limit; most children complete the test in well under an hour, making it a flexible, cost-effective and labour-saving screening device (Raven et al., 1990).
6. Used in conjunction with an equally valid test of verbal intelligence, the Raven's CPM can thus function equally well as a quick, effective screening instrument or as an important component of a more extensive clinical test battery.
7. Finally, it is generally agreed that for the test to accurately assess non-verbal intelligence in populations other than those for which it was developed, local norms are considered best (Sigmon, 1983). Thus, it is argued here that the development of

more appropriate normative data for the Raven's CPM for the population under study will make a significant contribution towards overcoming the problems currently facing South African psychologists (in both clinical and research domains).

It is argued further that the legacy of past inequalities inherent in primary and secondary schools in peri-urban and rural township settings have resulted in 'disadvantaged' populations that are vastly different to those on which the test was originally standardised overseas. Vass (1992) notes that the term 'disadvantage' is a relatively new one in psychology, particularly in the South African context. In this discussion, Vass' definition of the term is adopted wherein disadvantage is defined as referring to "the unequal access of students to education as a result of the oppression of black people in South Africa" (p. 106).

To conclude, although used extensively (usually as a screening instrument) across a wide variety of settings in South Africa, there are currently no available South African norms for the Raven's CPM. This pilot study therefore seeks to make a positive contribution to the effective channeling of Xhosa-speaking primary school children in peri-urban or 'township' educational settings, by developing much needed local normative data for the Grahamstown region.

Chapter Three

Methodology

The research method adopted in this study is that of a descriptive, normative study wherein the main findings will constitute a set of local normative data for the Raven's CPM for the target population under study. This methodological approach further limits the applicability of the normative database generated here in that it may only be used in those instances when the Raven's CPM is administered according to the modified method of test presentation adopted in this study.

3.1 Participants

Sample selection was carried out in consultation with staff at the Grahamstown Education Support Centre (a subsidiary of the Department of Education). The principle governing selection was the approximation (as closely as possible within the constraints of the limited funding and available human resources) of the population of young Xhosa-speaking primary school children currently being assessed at outpatient clinics in the township areas surrounding Grahamstown.

A Public Primary School was selected as being the most representative example of a Xhosa-medium Public Primary School in this area. The sample was drawn from a population of 644 Xhosa-speaking pupils currently attending this Primary School, which is located in Joza township on the outskirts of Grahamstown. Permission was obtained from the Department of Education and from the administrative body of the school itself [6]. This was granted on the condition that testing did not interfere with

[6] See Appendix 1.

the midyear scholastic achievement examinations, and would take place only once these had been completed.

The decision to test pupils by class (as opposed to individually) was forced on the researcher by the lack of available human and funding resources necessary for the assessment of large samples of individuals. As a result, the Grade 1 pupils were considered too young to be administered the test in the group format, and therefore had to be excluded from the sample. Each grade was comprised of two classes and fortunately the allocation of pupils to classes is random, with no distinction being made on the basis of academic ability. All pupils present in class on the day of testing were included in the sample.

3.2. The Instrument

The Raven's CPM test is a non-verbal test of intellectual ability and is regarded as being relatively free of accumulated knowledge. According to Raven et al. (1990), the Raven's CPM test gives an indication of the level of analogical thinking and abstract thought that a person has achieved and for this reason, it is a good measure of the level of intellectual development that s/he is able to utilise at this stage. In this study, the booklet form of the test was used. For a more detailed discussion of this instrument, see Chapter Two.

3.2.1 Visual Aids

Two sets of visual aids were constructed for the purposes of administering this test. The first consisted of a reproduction of the record form (or answer sheet) that had

been enlarged to three times the size of the original. The record form requested the following demographic data: (1) Name, (2) Age, (3) Grade, (4) Number of grades failed and finally, (5) a check box to indicate whether or not each participant had attended a preschool or a bridging class before beginning Grade 1. The second set of visual aids comprised reproductions of the first two items of the Raven's CPM test (to be used as "practice examples") that had also been enlarged to three times the size of the originals. These aids were constructed in such a way that the six options below the Matrices of Item 1 and 2 (from which the participant is expected to select the correct answer) were represented in the form of movable parts. This modification is recommended by Vass (1992) and functions to facilitate the physical placement of each of the six options within the matrix itself during the instructional phase of the test, thereby allowing the respondents to actively participate in the instructional phase.

It must be noted that the visual aids utilised in this study differ slightly from those recommended by Raven et al. (1990), where only the record form and Item 1 are reproduced, and enlarged to just twice the size of the original. As mentioned in Chapter Two, these modifications (and those made to the instructions below) were based on the work of Schwarz (1963), Crawford-Nutt (1976), Pons (1974) (cited in Vass, 1992) and more recently, that of Vass (1992).

3.2.2 The Translated Instructions

As the instructions for both the Raven's CPM and SPM are identical, the Xhosa instructions for this test as translated and adapted by Vass (1992) were also used in this study. These instructions had not only been translated into Xhosa (using the back translation method of Brislin, 1980) but had also been extended to cover the

inclusion of Item 2 as an additional practice example. Although the same instructions were used, a few minor spelling mistakes were observed in the original instructions as provided by Vass and these were corrected for this study[7].

3.3 Procedure

3.3.1 Data Collection

The approach towards data collection also followed Vass' (1992) administrative procedure and recommendations closely. The tester was fluent in Xhosa and introduced herself and the test to the class as follows:

(1). It was explained that it was important for psychologists at Rhodes University to obtain a group score on the test they were about to complete. The participants were assured that the test was not part of the school curriculum, and would in no way affect their existing scholastic achievement test results. It was further explained that every class (excepting those in Grade 1) in the school would be asked to complete the same test.

(2). The participants were asked whether there are any further questions, and these were addressed by the test administrator.

(3). Thereafter, the record form was unveiled and participants were asked to fill in their birthdate, gender, grade, number of grades failed, and whether or not they had attended a preschool or a bridging class before beginning Grade 1 (where necessary, the test administrator or the class teacher provided assistance with this task).

[7] See Appendix 2 for a copy of this set of instructions.

(4). The respondents were then shown the first (and later the second) test item and asked to actively participate by giving their answers. In each case, the movable part corresponding to their answer was physically placed in the test item's matrix and it's suitability discussed.

(5) The participants were informed that there was no time limit and were instructed to raise their hands as soon as they had finished, when either the test administrator or the class teacher made a note of the time taken to complete the test.

(6) Finally, at the end of each testing session, the record forms were placed in an envelope on which was written: the grade; the class number; the number of pupils present at school on the day of testing; the number of pupils assigned to each class; the class teacher's name and finally, the time the test was started.

3.3.2 Data Processing and Analysis

The demographic information collected from the participants was checked by their class teacher and verified by the author against the School's registration records, whereafter any necessary corrections were made.

The tests were scored out of a maximum score of 36 and statistics such as the means, standard deviations, range and frequency distributions of the scores were generated and calculated for each age group (at half yearly intervals). The participants' performance was compared across the following variables: age, gender and grade, preschool attendance and number of failed grades. The age related norms were also compared with those generated by Raven et al. (1990) for participants between the ages of 7 and 11 years 6 months. Finally, all descriptive and inferential statistical

procedures were conducted using the Statistica computer programme and these are outlined in detail in the following chapter.

Chapter Four

Results

In this chapter, the results concerning the following findings are outlined:

(1) Whether (based on the findings of Notcutt [1949] and Vass [1992]) the norms obtained are lower than those obtained by Raven et al. (1990).

(2) Whether the argument for developing more appropriate local normative data for the Raven's CPM in populations other than those for which the test has been standardised is supported by the findings of this study.

(3) Whether the test scores obtained for the Raven's CPM increase with age as argued by Raven et al. (1990).

(4) And finally whether, according to Vass (1992) and contrary to the findings of Raven et al. (1990), the scores obtained for this test are significantly influenced by education and gender (with males in this sample scoring significantly higher than females).

This outline begins with the demographic data regarding the normative sample and is followed by the results of the investigation of the impact of the relevant independent variables on the test scores obtained. Finally the normative data obtained for each age group is presented at half-yearly intervals. For a more detailed discussion of the implications of these findings in the context of the current study's aims and hypotheses, see Chapter 5.

4.1 Description of the Sample

Due to the fact that the test was administered in the week after the school examinations had ended, a regrettably high rate of absenteeism was evidenced. The final sample therefore consisted of just 380 Xhosa-speaking children. With regard to gender, 198 participants were male and 182 were female. More than half of the participants (57.4%) reported having undergone one or more extra years of schooling (i.e. have attended at least one year at a pre-school or a bridging year before starting Grade 1). Of the total sample, 117 participants (30%) had failed once; 45 (11.8%) had failed twice; 10 (2.6 %) had failed three times (2.6 %) and finally, 3 (0.8%) participants had failed four times. Table 1. (p. 46) presents the demographic data for this sample as it was originally collected (by grade). The highest number of participants fell into Grade Three where a total of 84 children were tested.

4.2 Breakdown in terms of Age and Gender

The sample was subdivided into half-yearly age intervals, according to the format in which the Raven's CPM normative data is traditionally presented. The breakdown of the sample in terms of age and gender is presented in Table 2. (p. 48).

The upper and lower limits of age for the total normative sample ranged from 6 years 7 months to 20 years 4 months. The latter participant was the only one to be excluded from the sample, as he was found to be significantly older than his peers. This extensive age range found in this sample highlights the fact that on average, most of the primary schools in these regions have an older population than would normally be expected.

Table 1: Sample - by GRADE

GRADE	n	MEAN AGE	AGE RANGE	GENDER		FAILED GRADES (1 yr or more)	PRESCHOOL (1 yr or more)
				M	F		
Grade 2 (Sub B)	63	8.26	6 – 12 yrs	31	32	18	59
Grade 3 (Std 1)	60	9.56	7 – 13 yrs	38	22	25	39
Grade 4 (Std 2)	45	10.26	8 – 17 yrs	22	23	11	29
Grade 5 (Std 3)	84	11.85	9 – 16 yrs	54	30	26	35
Grade 6 (Std 4)	61	13.21	10 – 20 yrs	25	36	44	29
Grade 7 (Std 5)	67	13.94	11 – 17 yrs	28	39	51	27

In addition, due to the low number of participants in the extreme lower and upper age ranges of the sample, the decision was made to collapse the scores of participants in those categories into the 7 and 16 year 6 months age groups respectively. It is thus important to emphasise that the latter two age groups are *mixed* groups.

4.2.1 The Mean Ages of the Participants.

The largest number of participants fell into the 11 year old age group, while the mean age of participants in this sample was 11. 2 months. This highlights the fact that on average, most of the primary schools in this region have an older population than would normally be expected. It must be noted that this finding has serious implications for this study, in that the Raven's CPM was specifically designed for children between the ages of 5 years 6 months and 11 years 6 months. Raven et al. (1990) assert that (in western samples) eductive ability level has usually been developed to the individual's full potential by the age of 11 years 6 months. Older children are therefore usually asked to complete the SPM version of the Raven's Matrices, which is deemed more appropriate for that age group.

The decision was therefore made to further subdivide the normative sample into two sets of norms. The first set is for children between the ages of 7 years and 11 years 6 months, while the second set applies to those children aged between 11 years 6 months and 16 years 6 months. This modification allowed for the qualitative comparison of the first set of normative data with those generated by Raven et al. (1990) for children in the same age groups.

Table 2: Sample - by AGE GROUPS

AGE	N	GENDER		AGE GROUP CATEGORY
		M	F	
6.5	2	1	1	6.03 – 6.08
7	2	1	1	6.09 – 7.02
7.5	6	3	3	7.03 – 7.08
8	14	5	9	7.09 – 8.02
8.5	19	12	7	8.03 – 8.08
9	27	13	14	8.09 – 9.02
9.5	16	5	11	9.03 – 9.08
10	27	10	17	9.09 – 10.02
10.5	31	19	12	10.03 – 10.08
11	32	16	16	10.09 – 11.02
11.5	24	13	11	11.03 – 11.08
12	18	13	5	11.09 – 12.02
12.5	23	11	13	12.03 – 12.08
13	22	9	13	12.09 – 13.02
13.5	21	11	10	13.03 – 13.08
14	27	15	12	13.09 – 14.02
14.5	14	6	8	14.03 – 14.08
15	22	12	10	14.09 – 15.02
15.5	14	10	4	15.03 –15.08
16	5	2	3	15.09 –16.02
16.5	6	4	2	16.03 – 16.08
17	5	4	1	16.09 – 17.02
17.5	2	2	0	17.03 – 17.08
20.5	1	1	0	20.03 – 20.08

With regard to age and gender, the mean age of female participants in the sample was 11 years, while the mean age of male participants was slightly higher at 11 years and 5 months. The result of a t-test comparing the differences in ages between the genders is presented in Table 3 below.

Table 3: t-test Results for Comparison of Differences in Ages of the Genders

Gender	t-value	Degrees of Freedom	Significance
Male/Female	1.615	377	0.107026

The results revealed no significant difference in age between the genders ($p > 0.05$).

4.3 The Scores on the Raven's CPM: Normative Sample

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.

4.3.1 The Distribution of the Scores

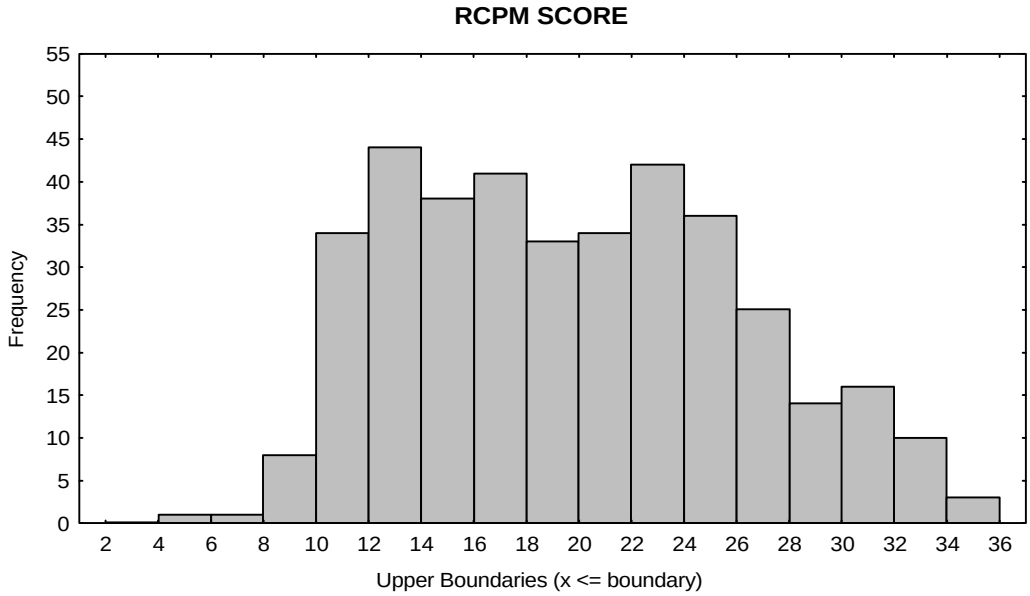
The distribution of the scores for the normative sample presented in the form of a frequency distribution in Figure 1 and Figure 2 on page 50.

.

Figure 1: Frequency Distribution of Scores on the RCPM

Score	*f*	Cumulative Frequency	Cumulative Percentage
0 – 4	0	0	0
5 – 8	2	2	0.5
9 – 12	42	44	11.6
13 – 16	82	126	33.2
17 – 20	74	200	52.6
21 – 24	76	276	72.6
25 – 28	61	337	88.7
29 – 32	30	367	96.6
33 – 36	13	380	100.0

Figure 2: Graph of Frequency Distribution of Scores on the RCPM



The distribution reveals a normal distribution. There is a sharp increase in the frequency of scores from the lowest score of 6 to the maximum frequency at a score of 14. Thereafter follows a sharp decline to a score of 20 (the mean of the sample) and finally, the distribution ends with a gradual decline in frequency to the highest score of 35.

4.3.2. The Mean, Standard Deviation and Range of Scores

The mean score, standard deviation, range, minimum and maximum scores for the Sample are presented in Figure 3 below.

Figure 3: Mean, Standard Deviation, Range, Maximum and Minimum Scores

Mean	Standard Deviation	Range	Minimum Score	Maximum Score
20.24	6.35	29	6	35

A mean score of 20.2 was obtained from 377 participants in the normative sample, while the standard deviation was 6.35.

4.4 Comparison of Scores across the Independent Variables

The participants' scores were analysed in relation to independent variables as recorded in the demographic data section of the record form.

4.4.1 Comparison of Scores Across Preschool & Failed Grades

The total raw test scores were correlated with whether or not the participants had attended pre-school or a bridging year before starting Primary School. The results showed that those who had not attended pre-school in fact performed better than those who had. However, this finding can also be explained in terms of the fact that pre-school education for children in these disadvantaged areas is a relatively new phenomenon. This explanation is supported by the fact that it was mainly the older children in the sample who had not experienced pre-school, and these participants performed better as a result of their age. No significant relationships were found between the number of failed grades and the participants performance on the Raven's CPM.

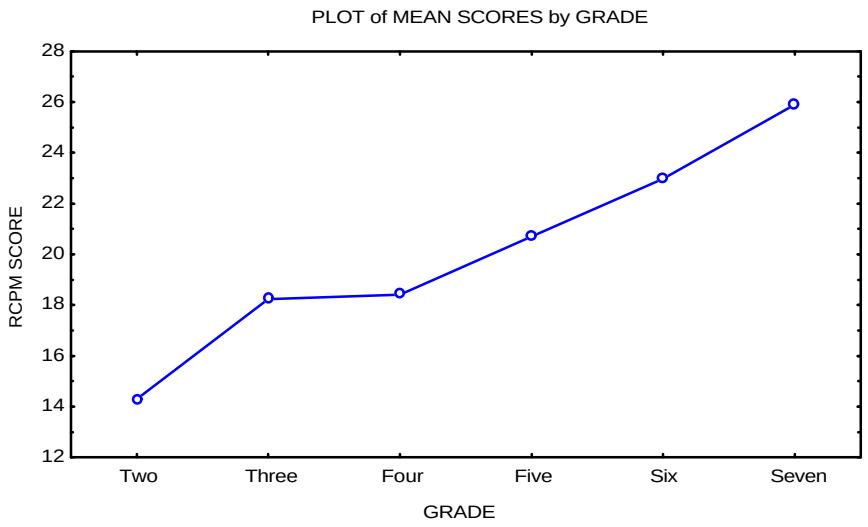
4.4.2 Comparison of Scores across Grades

In this section the raw scores obtained by the sample on the Raven's CPM are compared across the grades. Although the normative data is ultimately to be presented by age group and in the form of a percentile rank, the mean scores and standard deviations for each of the grades were generated and have been presented in Table 4 (p. 53). There is a steady increase in the mean CPM scores for each grade with the exception of Grade 3 and 4, where there is very little difference between the means. This data is also presented in the form of a cell mean plot in Figure 4 on page 54.

Table 4: Means, Standard Deviations, Range and Minimum and Maximum Scores by GRADE

GRADE	N	MEAN	STANDARD DEVIATION	RANGE	MINIMUM SCORE	MAXIMUM SCORE
Two	63	14.3	3.59	17	9	26
Three	60	18.2	5.57	26	6	32
Four	45	18.4	5.62	22	9	31
Five	84	20.7	4.90	21	11	32
Six	60	22.8	5.38	22	11	33
Seven	6	25.9	5.87	22	13	35

Figure 4: Plot of mean Raven's CPM scores by GRADE



A one-way ANOVA comparison of the mean Raven's CPM scores by grade was found to be significant at the 1% level. Post-hoc comparisons of means using Tukey's HSD for unequal sample sizes revealed significant differences between all grades except 3 & 4; 3 & 5; 4 & 5 and 5 & 6.

It is however important to note that**,** the age distribution within each standard is inconsistent and in most of the grades the participants' ages ranged across six or more years. First world countries usually have a set age range for each of the grades and children are expected to enter and graduate from primary schools at specific ages (usually 6 years and 12 years respectively). Here (as is the case with most of the schools in the township) the situation reflects children, who for various economic and socio-political reasons have started school later, or have been forced to miss school periodically (sometimes missing up to two years at a time before resuming their education). In some cases, the day, month and/or year in which the child was born is unknown, due to high rates of illiteracy among parents and relatives (personal communication, Mrs T. Hlengane, Educational Support Centre). Thus, the normative

data supplied by grade should only be used qualitatively, and the age related norms used instead for formal assessments. The results of the post-hoc comparison of means between Grade Two and Grades Three, Four, Five, Six and Seven indicate a significant difference between Grade Two and all the other Grades. The results of the post-hoc comparison of means between all the grades have also been attached (see Appendix 3).

4.4.3 Comparison of Scores across Gender

The mean score for the male participants in the sample was 21.3, with a standard deviation of 6.35, while the mean for the female participants was 2.2 points lower at 19.1 with a standard deviation of 6.18. Both genders attained a maximum score of 35, while the minimum score for female participants was one point lower than that of the males, who attained a minimum score of 7. A t-test was conducted to check whether the difference in means scores between the genders is significant and the results are presented below in Table 5.

Table 5: t-test Results for the Comparison of scores across GENDER

Gender	t-value	Degrees of Freedom	Significance
Mal e/Female	3.3030	377	0.0010

The t-test results reveal a difference between the means scores of the genders for the whole sample which is significant at the 1 per cent level. Males thus scored significantly higher than females in this normative sample. The results of independent t-tests for each pair (genders) at each of the age groups sampled have been attached (see Appendix 4). The findings showed that for many of the groups there is no a significant difference but, nonetheless the males do seem to consistently

outperform the females. A trend wherein boys consistently performed better was thus evidenced, however it is important to note that small group size precludes any categorical statement or claim to this effect. The decision was therefore made to refrain from separating the normative data generated in this study into two different sets of norms (one for each of the genders).

4.4.4 Comparison of Scores across Age

The mean scores for the sample by age are presented in Table 6 (p. 57). As mentioned earlier, only one participant's score was excluded from the sample as he was aged 20 years 4 months and thus significantly older than his peers. Also, due to low number of participants in the extreme lower and upper age ranges of the sample resulted in the decision to collapse these participants' scores into the 7 year old and 16 years 6 months age groups respectively. These are therefore *mixed* groups. A gradual increase in scores over age is evidenced; however due to idiosyncrasies in this normative sample (such as a maximum score of 22 at the young age of 8 years 6 months), these norms required smoothing.

4.5 The Norms

The smoothed norms for all the half-yearly interval age categories are presented in the form of percentile ranks in Figure 5 (7 years and 11 ½ years) and Figure 6 (11 ½ years to 16 years) (p. 58)[8] When the norms for participants between the ages of 7 and 11 years 6 months are compared with those of Raven et al. (1990), they appear to be lower than those obtained during the standardisation of this instrument in the United Kingdom and America.

[8] The unsmoothed/raw normative data has been attached, see Appendix 6.

Table 6: Mean Scores, Standard Deviation, Range, Minimum & Maximum Scores by AGE

AGE	n	MEAN	SD	MIN	MAX	RANGE
6.5	2	12	1.41	11	13	2
7	2	13	1.44	12	14	2
7.5	6	13.3	2.50	10	17	7
8	14	14.6	3.15	11	23	12
8.5	19	17.4	6.55	10	32	22
9	27	14.7	3.10	7	24	17
9.5	16	15.1	5.41	11	21	10
10	27	16.8	5.52	6	28	22
10.5	31	19.8	5.13	10	28	18
11	32	20.0	5.84	10	32	22
11.5	24	22.1	5.73	11	34	23
12	18	20.9	5.27	12	33	21
12.5	23	22.1	7.47	14	33	19
13	22	23.8	7.09	10	35	25
13.5	21	22.1	7.47	9	35	26
14	27	23.7	5.26	12	33	21
14.5	14	22.7	5.28	12	32	20
15	22	24.0	5.98	11	33	22
15.5	14	23.9	5.47	15	31	16
16	5	23.8	6.87	16	32	16
16.5	6	23.8	5.75	18	33	15
17	5	21.4	6.39	13	29	16
17.5	2	25	9.90	18	32	14
20.5	1	30	-	?	?	?

Finally, a relatively high Cronbach Alpha coefficient of .88 was found for this test indicating that despite the problematic small sample sizes, the Raven's CPM test proved to be a reliable measure of nonverbal intelligence within this population.

Figure 5. Smoothed Norms for the Raven's CPM for the Lower Age Range

PERCENTILE POINTS	CHRONOLOGICAL AGE IN YEARS <7	7.5	8	8.5	9	9.5	10	10.5	11	11.5
95	14	17	19	20	21	21	24	27	28	30
90	14	16	18	19	20	20	23	25	26	28
75	13.5	15	16	16	17	17	18	21	23	24
50	12.5	13	13.5	14	14	14.5	16	18	20	20
25	11.5	12	12	12	13	13	13	14	15.5	16
10	10	11	12	12	12	12	13	13	14	14
5	9	10	11	11	11	12	12	12	13	13

Figure 6. Smoothed Norms for the Raven's CPM for the Upper Age Range

PERCENTILE POINTS	CHRONOLOGICAL AGE IN YEARS 12	12.5	13	13.5	14	14.5	15	15.5	16	16.5
95	31	31	32	32	33	33	34	34	35	35
90	28	29	30	30	31	31	32	32	33	34
75	24	25	25	26	26	27	27	28	29	30
50	21	22	22.5	23	23	23	24	24	24	25
25	16	16	17	17	18	18	19	19	21`	20
10	14	15	15	16	16	16	16	17	17	17
5	13	14	14	14	15	15	15	16	16	17

.

Chapter Five

Discussion and Conclusion

5.1 Discussion of Results

In this section, the results obtained will be discussed in relation to: (1) the South African context (2) the relevant literature, (3) the limitations of the current study; and wherever possible, (4) the findings obtained by Raven et al. (1990).

The principal findings of this study can be summarised as follows:

(1) A significant effect of age on test scores was evidenced, where scores increased with age as expected.

(2) A significant effect of gender on test scores was also found, with males performing significantly better than female participants. However, the small group size precludes any categorical claim to this effect, and a consistent tendency for males to perform better than females is merely noted, as is the need for further research in this area.

(3) Due to the vast age range (with age gaps of up to six and ten years) in each of the grades sampled, the results of one-way analysis of variance between the grades also proved to be inconclusive.

(4) No significant effect of education was observed on test scores in that no relationships were found between the number of failed grades and/or attendance at pre-school (or a bridging year) and performance on this test.

(5) And finally, qualitative comparison indicates that the norms generated for this sample generally appear lower than those obtained by Raven et al. (1990) during the standardisation of this instrument in the United Kingdom and America.

5.1.1 The Effects of Age on Test Scores

The mean scores are presented by age in Table 6 (p. 57). The trend is for the Raven's CPM scores to increase with age up to the age of 11 years 6 months, with the exception of the 8 years 6 months age group (where an exceptionally high score of 32 elevated the mean to 17.4). This finding supports Raven et al.'s (1990) finding that test scores on the Raven's CPM increase with age and reach their 'ceiling' by the age of 11 years 6 months.

Thereafter, an inconsistent pattern in the mean test scores for participants in the 13.5 and older age groups is evident and this is most likely to be a reflection of the effects of a "self-selection process" (Vass, 1992, p. 94). The latter refers to the phenomenon where less intelligent children in Primary and Secondary schools in disadvantaged rural and peri-urban township settings are generally allowed to remain at school indefinitely, despite the fact that they are significantly older than their peers and making very little progress through the curriculum. This practice is best understood as being a residual effect of the previous Government's separate development policies, the 'Bantu Education System' in particular (Vass, 1992), and the resultant lack of special education facilities in these disadvantaged areas. In contrast, it is highly unusual in western populations (such as those in which the test was originally developed) that any child who is significantly older than his or her peers would be allowed to remain in the mainstream curriculum for an indefinite period of time. In

addition, the oldest participant in this sample was over 20 years old and therefore significantly older than the next oldest participant (aged 17 years 7 months). For this reason, the decision was made to exclude his score from the sample.

Furthermore, as result of the unusually high rate of absenteeism experienced (due to the test being administered after the mid-year scholastic examinations had been completed) the sample sizes for most of the age groups in the sample were also relatively small. The numbers of participants in the extreme upper and lower age ranges were particularly low and resulted in the decision to collapse these participants' scores into the 7 year old and 16 years 6 months age groups respectively. Finally, the above-mentioned factors also necessitated the separation of the normative data into the upper (11 years 6 months to 16 years 6 months) and lower (7 years to 11 years 6 months) age ranges and a smoothing of the age-related norms. This data is presented in Figures 5 & 6 (p. 58).

5.1.2 The Effect of Gender on Test Scores.

Despite the fact that there was no significant difference in age between the genders, a significant difference in performance (at the one percent level) was evidenced, with males scoring significantly higher than the females. The results of the t-test for the whole group are presented in Table 3 (p.49), while those for each of the age groups have been attached in Appendix 3. However, as noted earlier, the problem of small group size (also due to the high rates of absenteeism experienced) precludes any categorical statement to this effect, and instead a consistent trend in which males outperformed females is merely noted. This finding is contrary to that of Raven et al.

(1990), who found no significant difference between the genders with regard to performance on the Raven's CPM.

It is however, supported by Vass (1992), who found significant gender differences (with males scoring higher than females) in performance on the Raven's SPM, in a similar but older sample of children. Thus, it appears that when assessing non-verbal intelligence in impoverished and disadvantaged populations such as these, the difference in the scores between the genders may need to be taken into consideration. In light of the above findings and their implications, it is therefore strongly recommended that further research in this area be undertaken in the future, in order to either confirm or refute the need for separate sets of normative data for each of the genders on this test.

5.1.3 The Effect of Grade on Test Scores

A steady increase in the mean scores for each grade was evidenced, with the exception of Grade 3 and 4 (where there was very little difference in the mean scores). The results of a one-way analysis of variance revealed a significant difference between Grade 2 and all the other standards. The results of post-hoc comparison of means using Tukey's HSD for unequal sample sizes revealed significant differences between all grades except 3 & 4, 3 & 5, 4 & 5 and 5 & 6 (see Appendix 4). Although this normative data is presented by grade in Table 4 (p. 53), the unusually wide range of ages evidenced within each of the grades (with the ages in most grades spanning more than six years) indicates that the more reliable age-related normative data should be used instead.

5.1.4 The Effect of Failed Grades and Preschool Attendance on Test Scores

No significant relationships were found between the participants' scores on the Raven's CPM and the number of grades failed. In the case of preschool attendance, the results showed that participants who had not attended preschool actually performed better than those who had. This finding may be reflecting the fact that, for most of the residents in impoverished township settings throughout South Africa, the option of sending their children to pre-school (as opposed to a crèche or child-minder) is a relatively new one. With regard to the population under study, the majority of the older children in this area did not have the advantage of having had a preschool education. Thus, it is possible that this finding may be a function of the effects of the expected influence of age on the test scores. In addition, the above findings offer further support for the argument that scores derived from this test are not influenced by education (Raven et al., 1990).

5.2 The Norms

For reasons outlined in Chapter Four, the age-related norms required smoothing and are presented in Figures 5 & 6 (p.58). Despite the problems experienced with the extended range of ages and small sample sizes, a Cronbach Alpha Coefficient of .88 for the whole range of development was obtained. This result suggests that, when these norms are applied appropriately (i.e. for Xhosa-speaking Primary school children in the disadvantaged educational settings in the townships surrounding

Grahamstown); the Raven's CPM test will produce a valid and reliable assessment of non-verbal intelligence. In addition, the merits of separating the normative data into two sets of norms (due to the extended age range within this sample) require further discussion. It is argued here that the norms for the lower age range (i.e. 7 years to 11 years and 6 months) are valid and generally considered to be unproblematic, but that the older set of norms represents somewhat of a dilemma. In theory, children in this age group should have been administered the Raven's SPM version of the test. However, the fact that this particular group has been unable to progress through the Primary school curriculum raises the question of whether CPM norms are not perhaps more appropriate in this context.

It is argued here that it is precisely because these children are experiencing considerable difficulty in their scholastic development, that the Raven's CPM would best be able to identify their current level of non-verbal intelligence and thus hopefully provide some indication as to the nature of their difficulties. For this reason, the second set of data has been included here and the decision regarding which of the two versions of this test to administer has been left up to the individual.

5.2.1 Comparison of Norms with Raven et al. (1990).

It was found that despite the improved method of test administration recommended by Schwarz (1963), Pons (1974), Crawford-Nutt (1976) and Vass (1992), the norms obtained appear qualitatively lower than those obtained by Raven et. al (1990). These findings are supported by those of (Vass, 1992), who found significantly lower norms for the Raven's SPM when administered to a similar but older population of disadvantaged children. As was the case with Vass' study, in this instance the mean

score of most of the age groups is at least ten points lower than those obtained by Raven et. al (1990).

This finding can be understood using Vass' definition of the concept of disadvantage as discussed in Chapter Two. The impoverished conditions in which these children live (and attend school) are likely to have had a negative effect on their intellectual development. One example is that most of the children have not eaten before going to school and some receive their only meal of the day at school, in the form of soup kitchens set up by the government's department of Education and various international aid agencies (personal communication, School Principal of the Primary School under study). The effects of malnutrition on cognitive development have been widely documented both internationally, and in South Africa (Griesel & Richterand, 1987), and it is likely that high rates of prevalence also exist in the target population under study.

Other less obvious but equally insidious factors may also be responsible for the lower norms. Due to socio-economic problems (such as geographic marginalisation, dependence on 'unsafe' public transport, crowded living conditions, congested roads and a lack of safe and adequate play ground facilities), high rates of closed head injuries are also found to exist among children living in these areas (Jacobs, 1984). In addition, generally high levels of physical abuse and violence further increase the prevalence rates for this condition. Conditions in the schools themselves may also have affected the participants' levels of concentration and thus contributed to their lowered performance on this test. In this study, the test was administered during winter, in classrooms that had many broken windows and no electricity. In addition,

many of the participants were not warmly dressed and a significant number were not wearing shoes. It is possible that the low temperatures and inadequate lighting may have impaired their ability to concentrate, as their main concern was keeping warm.

A different yet complementary explanation for these results is related to the cross-cultural research investigating the effect of culture on childhood growth and development. As mentioned earlier, Serpell (1994) investigated the lowered performance of African children (particularly those living in rural areas) on tests similar to the Raven's CPM (ie. the Block Design assembly test, the Bender Gestalt test and the Draw-a-person test) in comparison to their western peers. His findings revealed that the "abstract psychological function of pattern reproduction manifests itself according to the demands of the eco-cultural niche to which the subjects behaviour is adapted" (1979, cited in Lonner & Malpass, 1994, p. 159). It is thus possible that the lowered test performance is reflecting more the lack in their environment of the necessary materials (such as paper and pencils) and the resulting lack of 'practice' or experience needed to accomplish this type of pattern reproduction task/activity successfully, rather than any deficit in non-verbal intellectual ability itself.

Thus to conclude, the lower norms are understood as reflecting both the effects of living under impoverished conditions, the participants' disadvantaged backgrounds and environments and the forms in which pattern reproduction skills are both manifested and valued within this particular eco-cultural context, and not that black or Xhosa-speaking children are inherently less intelligent than their white peers. In addition to emphasizing the need for more appropriate local norms for the Raven's

CPM, it is this author's intention to counteract the above ill-informed and extremely simplistic interpretation of the same data that was so characteristic of the South African research during the Apartheid era.

5.3 Limitations of the Study and Recommendations for Further Research

In this section, the limitations of the current study are outlined and discussed in detail. Wherever possible, suggestions or recommendations for further research are made; which if followed, will serve to extend to severely limited database currently existing for this test in South Africa.

(1) The choice of opting for an 'opportunistic' sample selection procedure (i.e. testing everyone who is currently attending this Primary school) proved to be problematic for two reasons. Firstly, no control existed over the numbers of children in each of the relevant age groups. Thus, small sample sizes were experienced, particularly with regard to the younger age groups. Secondly, it was only once the data had been collected that the true age range of the sample could be known. In other words, it was only once the testing had been completed that it was discovered that just under half of the sample were older than 11 years 6 months, and (theoretically) should have been administered the Raven's SPM. Thus, the efficacy of the normative data collected in this study was considerably undermined by the problem of small sample sizes.

(2) Secondly, the timing of the data collection also proved to be problematic. Permission to test at the school was granted on the condition that it was administered after the mid-year scholastic examinations had been completed. In

the planning stages, the author did not consider that this would lead to high rates of absenteeism, and ultimately to small sample sizes. In the future research, it is recommended that arrangements be made to ensure that the test is administered during a time period when most of the children are expected to be in attendance.

(3) Furthermore, although great care and consideration was taken in selecting a Primary school that was as representative of the target population as possible, the fact remains that only one school was sampled. The latter was determined by the availability of funding, human resources and the scope and aims of the current study. Thus, although both useful and valid, the normative data is nevertheless limited to a very circumscribed Xhosa-speaking population in the peri-urban township areas surrounding Grahamstown. Recommendations for further research therefore also include the importance of obtaining normative data for Xhosa-speaking children attending private schools and academies. It is likely that, as a result of having a more privileged background, these children will perform better than those in the current study on the test, and this matter requires further investigation.

(4) The children in Grade One also had to be excluded from the sample as they were considered too young to be administered the test in group format. This is a serious limitation, as school readiness testing of Xhosa-speaking children is a major concern for parent's and teachers in Primary Schools throughout South Africa. Although testing children individually is a costly and labour intensive process, it is highly recommended that in future a study be conducted which focuses on developing norms specifically for children between the ages of 5 years and 7 years. It is also suggested that the scores obtained on this study be

correlated with another appropriate assessment instrument such as the Draw- A-Person Test in order to build up a much-needed database of South African norms.

(5) Finally, the lowered performance of the participants on this test (when compared with their western peers) has highlighted the importance of further investigation into the cultural and socioeconomic influences on the nature of non-verbal intelligence in these disadvantaged African populations.

5.4 Conclusion

It has been argued here that the Raven's CPM is a reliable and valid instrument for the assessment of non-verbal intelligence in South African children from all cultural and language groups. It is further argued that the CPM not only functions as a quick, cost-effective and accurate screening instrument, but it is also a valuable component in more in-depth diagnostic test batteries. The results of this study revealed the urgent need for the development of more appropriate local normative data for this test, particularly when it is being administered in disadvantaged populations of children in rural and peri-urban township areas. Furthermore, although no significant effect of education was noted, a consistent tendency for males to score higher than the females was evidenced. These findings (in addition to the lowered performance in general) highlight the importance of further research into the effects of cultural and socio-economic factors as well as gender, on non-verbal intelligence (and on performance on this test in particular) in these populations. Finally, it is argued here that process of assessing and placing children within the new South African school curriculum (Curriculum 2005) will be significantly improved through the establishment of further more appropriate local normative data for this labour-saving screening instrument.

References

Anastasi, A. & Urbania, S. (1997*). Psychological Testing (7th ed.).* New Jersey: Prentice-Hall Inc.

Ardila, A. (1995). Cross-cultural neuropsychology. *Journal of Clinical and Experimental Psychology, 17*, 1, 143-150.

Avis, C.E. (1995). *WISC-R Coding Incidental Recall, Digits Forward, Backward and Supraspan Test Performance in Children Aged 6 and 7.* Masters Dissertation, Rhodes University, Grahamstown.

Barnabas, I.P., Kapur, M. & Rao, S. (1995). Norm development and reliability of Coloured Progressive Matrices Test. *Journal of Personality and Clinical Studies, 11* (1-2), 17 – 22.

Berlioz, L. (1955). *Etude des Progressive Matrices faite sur les Africans de Douala.* Bulletin du Centre des Etudes et Recherches Psychotechniques, Douala, Cameroons.

Bracken, B.A & Barona, A. (1991). State of the art procedures in translating, validating and using psycho- educational tests in cross-cultural assessment. *School Psychology International, 12* (1-2), 119-132.

Brislin, R.W. (1980). Translation and content analysis of oral and written material. In H.D. Triandis & J. W. Berry (Eds.), *Handbook of cross-cultural psychology (Vol. 2)* (pp. 389-444). Boston: Allyn & Bacon.

Brody, N. (1992). *Intelligence. (2nd ed.)*. Toronto: Academic Press, Inc.

Cantwell, Z.M. (1967). The performance of American pupils on the Coloured Progressive Matrices. *British Journal of Educational Psychology,37*, 389-390.

Cattell, R.B., & Butcher, H. j. (1968). *The prediction of achievement and creativity*. New-York: Bobbs- Merrill Company.

Cattell, R.B., & Cattell, A.K.S. (1963). *Culture fair intelligence test*. Champaign, IL: Institute for Personality and Ability Testing.

Carlson, J.S., & Jensen, C.M. (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's Progressive Matrices Test. *Psychological Review, 97*, 404-431.

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

Cronbach, L.J., & Drenth, P.J.D. (Eds.) (1972). *Mental tests and cultural adaption.* The Hague: Mouton.

Dyall, K. (1996). *The Bender Gestalt Test – an investigation into problems concerning it's administration and scoring and it's application to low educated adults.* Unpublished Master's Thesis, Rhodes University, Grahamstown.

Foxcroft, C.D. (1996). *Psychological Assessment of children: It's relevance and challenges facing us in the New South Africa.* Unpublished manuscript, University of Port Elizabeth, Port Elizabeth.

Freeman, F.S. (1966). *Theory and practice of psychological testing.* New York: Rinehart & Winston.

Frijda, N & Jahoda, G. (1966). On the scope and methods of cross-cultural research. *International Journal of Psychology, 1,* 109-127.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences.* New York: Basic Books.

Griesel, R. and Richter, L. (1987). Psychosocial studies of malnutrition in South Africa. In G. Bourne (ed.), *World Review of Nutrition and Dietetics.* Basel: Kager.

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

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

Irvine, S.H. (1963). Some practical and theoretical problems of general ability testing at the African Standard Six level in Southern Rhodesia. *Conference Proceedings,* East African Institute of Social Research.

Irvine, S.H. (1979). The place of factor analysis in cross-cultural methodology and it's contribution to cognitive theory. In L.H. Eckensberger, W.J. Lonner & Y.H. Poortinga (Eds.), *Cross-cultural contributions to psychology* (pp. 300-341). Amsterdam: Swets and Zeitlinger.

Jacobs, M. (1984). Childhood head injuries and poverty. *Conference Proceedings,* Second Carnegie Inquiry into Poverty and Development in Southern Africa.

Jaworska, A. & Szustrowa, T. (1993). Polish Standardisation of the RPM. *Polish Psychological Bulletin, 24* (4), 303- 307.

Jensen, A.R. (1974). How biased are culture-loaded tests ?. *Genetic Psychology Monographs, 90,* 185 – 244.

Jensen, A.R. (1980). *Bias in Mental Testing.* London: Methuen & Co, Ltd.

Jensen A.R. & Rowther, W.D., JR. An experimental analysis of learning abilities in culturally disadvantaged children. Final Report, U.S. Office of Economic Opportunity, Contract N0. OEO 2404, 1970, 1 - 181.

Jensen, A.R. (1998). *The g factor: the Science of Mental Ability*. Praeger: Connecticut.

Kail, R & Pellegrino, J.W. (1985). *Human Intelligence: Perspectives and Prospects*. New York: Freeman & Co.

Kirby, J.R. & Das, J.P. (1990). A cognitive approach to intelligence: Attention, coding and planning. *Canadian Psychology, 58,* 287 –300.

Klingelhofer, E.L. (1967). Performance of Tanzanian secondary school pupils on the Raven's Progressive Matrices Test. *Journal of Social Psychology, 72,* 205-215.

Labuschagne, A.S. (1995). *Perceptual tests for distinguishing difference in ability among natives of Central Nyasaland: an adaptation of Raven's Progressive Matrices*. Unpublished Master's Thesis, University of South Africa, Pretoria.

Lezak, M.D. (1995). *Neuropsychological Assessment* (3rd ed.). Oxford: Oxford University Press.

Lonner, W. (1990). An overview of cross-cultural testing and assessment. In R.W. Brislin (Ed.), *Cross-cultural research and methodology series: Vol. 13 Applied cross-cultural psychology* (pp. 56–76). Newbury Park, CA: Sage Publications.

Lonner, W. & Malpass, R. (1994). *Psychology and Culture.* Boston: Allyn and Bacon.

Malan, A. (1951). *'n Voorlopige herstandaardisasie van Raven se patroontoets*. Unpublished Master's Thesis, Stellenbosch University, Cape Town.

Maloney, M.P. & Ward, M.P. (1976). *Psychological Assessment: a conceptual approach*. New York: Oxford University Press.

Maree, I. (1994). *The connection between the performance on the South African Wechsler Intelligence Scale for Adults (SAWAIS) and the Raven's Coloured Progressive Matrices (CPM) among mentally impaired patients.* Unpublished master's thesis, University of Stellenbosch, Cape Town.

Martin, A.W. & Wiechers, J.E. (1954). Raven's Coloured Progressive Matrices and the Wechsler Intelligence Scale for Children. *Journal of Consulting Psychology, 2,* 143 – 144.

Merenda, P. (1994). Cross-cultural testing: Borrowing from one culture and applying it to another. In L.L. Loebner & U.P. Gielen (Eds.), *Cross-cultural topics in Psychology* (pp. 53 - 58). Westport, CT: Praeger Publishers.

Notcutt, B. (1949). The measurement of Zulu Intelligence. *Journal of Social Research (Pretoria), I*, 195 –206.

Ogunlade, J. (1978). The predictive validity of the Raven's Progressive Matrices with some Nigerian Children. *Educational and Psychological Measurement, 38*, 465 – 467.

Pons, A. L. (1974). *Administration of tests outside the cultures of their origin*. Paper presented at the 26th Annual Congress of the South African Psychological Association, Johannesburg.

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

Raven, J.C., Court, J.H. & Raven, J.C. (1990). *Manual for Raven's Progressive Matrices and Vocabulary Scales - Section 2: Coloured Progressive Matrices.* Oxford: Oxford Psychologists Press.

Raven, J. C., Court, J.H. & Raven, J.C. (1993). *Manual for Raven's Progressive Matrices and Vocabulary Scales – Section 1: General Overview*. Oxford: Oxford Psychologists Press.

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

Retief, A. (1988). *Method and Theory in cross-cultural psychological assessment*. Human Sciences Research Council. Pretoria: South Africa.

Rust, J.R. & Golombok, S. (1989). *Modern Psychometrics: The science of psychological assessment*. Routledge: London.

Sattler, J.M. (1982). *Assessment of Children's Intelligence and Special Abilities* (3rd ed.) Allyn & Bacon, Inc. Massachusetts.

Schwarz, P.A. (1963). Adapting tests to the cultural setting. Educational and Psychological Measurement. 34 (4), 673 – 686.

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

Serpell, R. (1994). The cultural construction of Intelligence. In W.J. Lonner & R.S. Malpass (Eds.), *Psychology and Culture*. Boston: Allyn and Bacon.

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

Silvey, J. Testing ability tests: Issues in the measurement of ability among African Schoolboys. Mimeographed paper, *East African Institute of Social Research Conference Proceedings,* 1963. I

Shindler, J. (1997). Education in South Africa: A statistical overview. In H.P Africa, H. C. Charles, A. Miller, & S. Sisulu (Eds.), *Education Africa Forum*. (pp. 142-147). Johannesburg: Education Africa.

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

Sternberg, R.J. (1987). Synopsis of a Triarchic Theory of Human Intelligence. In S.H. Irvine & S.E. Newstead (Eds.), *Intelligence and Cognition: Contemporary Frames of Reference,* (pp. 141-169). Dordrecht: Martinus Nijhoff Publishers.

Sternberg, R.J. & Detterman, D.K. (1986*). What is intelligence ? Contemporary viewpoints on it's nature and definition.* Norwood, NJ: Ablex.

Valencia, R. (1984). Reliability of the Raven Coloured Progressive Matrices for Anglo and Mexican-American children. *Psychology in the Schools, 21*, 49 – 52.

Van der Vijver, F.J.R. & Leung, K. (1995). *Methods and data analysis in cross-cultural research.* Handbook of cross-cultural psychology. Manuscript submitted for publication.

Van de Vijver, F. J.R. & Poortinga, Y. (1995). Towards an integrated analysis of bias in cross-cultural settings. *European Journal of Psychological Assessment, 13* (1), 29-37.

Vass, V. (1992). *Standardisation of Raven's Standard Progressive Matrices for Secondary School African pupils in the Grahamstown Region.* Unpublished Master's Thesis, Rhodes University, Grahamstown.

Viljoen, G, Tredoux, C & Levett. A (1994). Using the Bender Gestalt in South Africa – some normative data for Zulu-speaking Children. *South African Journal of Psychology, 20* (3), 144-151.

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

Wright, S.C., Taylor, D.M. & Ruggiero, K.M. (1996). Examining the potential for academic achievement among Inuit children: Comparisons on the Ravens Coloured Progressive Matrices. *Journal of Cross-cultural Psychology,* 27 (6), 733- 753.

Yau-Fai Ho, D. (1994). Introduction to Cross-Cultural Psychology. In L.L. Loebner & U.P. Gielen (Eds.), *Cross-cultural topics in Psychology* (pp. 1- 32). Westport, CT: Praeger Publishers.

Appendix 1

The District Manager
Dept of Education
Private Bag X1001
Grahamstown
6140
Tel:(046) 622 9310
10th March 1999

Head of Department
Rhodes Psychology Department
Rhodes University
Grahamstown
6140
Attention: Prof Chris Stones

Re: Miss N. Bass -Master's Thesis in Clinical Psychology - 1999

Dear Professor Stones,

This is to certify that the Department of Education in Grahamstown has received and approved the above-mentioned student's Research Proposal.

Permission is hereby granted for Miss Bass to administer the Raven's Coloured Matrices Test to a representative sample of Xhosa-speaking primary school children in the Joza area. This sample will be drawn from a primary school population, deemed by the Department as reflecting an accurate cross-section of this community. The assistance and full cooperation of the staff, both at the Department of Education and the school in question, is also hereby guaranteed. The latter is on the condition that Miss Bass undertakes to ensure that there is minimal disruption to the school's daily routine during testing.

Miss Bass has indicated to me that she wishes to administer the tests during the second term this year, and once the dates have been finalised, the Department undertakes to assist with the necessary arrangements in order for testing to proceed as planned.

Should you require any further information, please do not hesitate to contact me at the above address.

Yours sincerely

R. Mgolodela
District Manager

THE DISTRICT MANAGER
GRAHAMSTOWN DISTRICT OFFICE
PRIVATE BAG X1001
GRAHAMSTOWN, 6140.

Appendix 2

Adapted/ translated instructions in Xhosa

Thyilani iincwadi zenu kwiphepha lokuqala. Linje. Phezulu ephepheni kukho umbhalo othi "Set A", nakumaphepha enu eempendulo kukho le kholam ka "Set A". Lo ngu A1. Niyambona ukuba unjani na. Ngasentla apha yipateni enendawana engekhoyo. Ipisi nganye kwezi zilapha ngezantsi (yalatha kwipisi nganye nganye) imile ngokufanelekileyo ngqo ukuba ingene gingci kwesi sikhewu, kodwa ayizizo zonke ezigqibezela le pateni ngokupheleleyo. Unombolo 1 (yalatha ipisi yokuqala ubuye walathe kwipateni) asiyiyo tu elungele ukungena apha..(Khumbula ukususa ipisi nganye uyilinganise esikhewini). Unombolo 2 naye akalungi, uyangena esikhewini kodwa akayigqibezeli ipateni. Ngaba unjani yena u-6? (Bavumele baphendule, bachaze nezizathu). Ngeba yipateni efanelekileyo (bonisa ukuba ipateni ifana nqwa naleyo ingasentla) kodwa ayiphelelanga ngokuthe twatse. Bekani iminwe yenu kuleyo ilunge ngokupheleleyo. Ewe u-4 ngoyena yena. Ngoko ke impendulo ka -A1 ngu- 4. Bhalani ke u-4 apha ecaleni kukanombolo 1 kukholam A kumaphepha enu eempendulo. Thyilani ke ngoku kwiphepha elilandelayo.

Lo ngu - A2. Niyambona? Lendawo ingasentla yipateni enendawana engekhoyo. Ipisi nganye kwezi zingezantsi (yalatha ipisi enye emva kwenye) imile ngokufanelekileyo ngqo ukuba ingene gingci kwesi sikhewu. Kodwa ayizizo zonke eziyigqibezela ngokupheleleyo ipatenile. Ngaba unjani unombolo -1? (Bayeke baphendule bachaze nezizathu). Uyangena gingci kwisikhewu, kodwa akayiyo ipateni efanelekileyo. Bekani iminwe yenu kuleyo ilunge twatsa. Ewe unombolo -5 ngoyena ulungileyo. Ngoko ke impendulo ka-A2 ngu- 5. Bhala u-5 apha ecaleni kuka nombolo 2 kukholam A kwiphepha lakho leempendulo.

Kwiphepha ngalinye encwadini kukho ipateni enendawana engekhoyo. Kufuneka ke qho uzigqibele ngokwakho ukuba yeyiphi na ipisi, kwezi zingezantsi, efanelekileyo ukugqibezela ipateni engentla. Wakuyifumana leyo ifanelekileyo bhala phantsi inombolo yayo kwiphepha lakho leempendulo, ecaleni kwenombolo yepateni. Ungabhali kwincwadana le yakho. Zilula ekuqaleni kodwa ziye zisiba nobunzima njengokuba ubheka phambili. Akukho ngxaki. Ukuba niyiqonde kakuhle indlela ezihamba ngayo ezi zilula akuyi kuba nzima kangako ukuzenza nezi zilandelayo. Linga enye emva kwenye, ukusuka ekuqaleni ukuya esiphelweni sencwadana. Sebenza ngesakho isantya. Ungazitsibi ezinye. Ungabuye

uthyile kwiphepha osele uligqithile. Khangela ukuba ungachana zibe ngaphi na. Ungathatha ixesha elide kangangoko ufuna. Yenzani elandelayo.

Enyanisweni elona lifanelekileyo ngu nombolo -1. Qinisekani ke ukuba nibhale u - 1 ecaleni kukanombolo 3 kukholam A kumaphepha enu empendulo. Qhubani ke kanjalo ngokwenu nide niyokufikelela esiphelweni sencwadana. Ndiza kumana ndijikeleza ndikhangela ukuba niqhuba kakuhle na.

Appendix 3

Results of Post-hoc Comparisons Of Mean Scores Across Grades

GRADES	Two	Three	Four	Five	Six	Seven
Two m = 14.3	-	.00048	.00228	.00002	.00002	.00002
Three m = 18.2	.00048	-	**.99998**	**.09247**	.00003	.00002
Four m = 18.4	.00228	**.99998**	-	**.28922**	.00045	.00002
Five m = 20.7	.00002	**.09247**	**.28922**	-	**.15053**	.00002
Six m = 22.9	.00002	.00003	.00045	**.15053**	-	.02258
Seven M = 25.1	.00002	.00002	.00002	.00002	.00258	-

Appendix 4

Comparison of male and female Raven's CPM scores by age group

Age Group	**Males**		**Females**		***t*-score**	**df**	***p*-value**
	n	Mean RCPM score	n	Mean RCPM score			
≤7	2	12.50	2	12.50	N/A	-	-
7.5	3	11.33	3	15.33	-3.618	4	.02239*
8	5	14.80	9	14.55	.134	12	.89589
8.5	12	19.33	7	14.29	1.703	17	.10674
9	13	15.15	14	14.29	.6089	25	.54808
9.5	5	16.20	11	14.73	.8729	14	.39743
10	10	17.40	17	16.59	.3703	25	.71429
10.5	19	21.95	12	16.58	2.953	29	.00618*
11	16	20.50	16	19.68	.4424	30	.66136
11.5	13	21.46	11	23.00	-.6343	22	.53245
12	13	21.15	5	20.40	.2428	16	.81123
12.5	11	23.36	12	21.00	.8112	21	.29291
13	9	23.56	13	24.08	-.1571	20	.87672
13.5	11	23.36	10	20.80	.8203	19	.42224
14	15	24.53	12	22.83	.8287	25	.41512
14.5	6	23.17	8	22.38	.2674	12	.79372
15	12	24.75	10	23.30	.5574	20	.58347
15.5	10	24.80	4	21.75	.9382	12	.36665
16	2	30.50	3	19.33	3.3855	3	.04292*
>16	11	24.00	3	21.33	.6482	12	.52904

**denotes significance*

Appendix 5.

Unsmoothed (Raw) Normative Data For Both Upper and Lower Age Ranges

PERCENTILE POINTS	CHRONOLOGICAL AGE IN YEARS									
	≤7	**7.5**	**8**	**8.5**	**9**	**9.5**	**10**	**10.5**	**11**	**11.5**
95	14	17	23	32	21	21	24	28	27	31
90	14	17	18	28	21	20	23	27	26	28
75	13.5	15	16	23	17	17	22	24	24.5	26
50	12.5	13	13.5	16	14	14.5	17	22	20	24
25	11.5	12	13	12	13	12	12	15	15.5	17.5
10	11	10	12	11	11	12	11	13	15	14
5	11	10	11	10	10	11	9	11	11	13

PERCENTILE POINTS	CHRONOLOGICAL AGE IN YEARS									
	12	**12.5**	**13**	**13.5**	**14**	**14.5**	**15**	**15.5**	**16**	**16.5**
95	33	31	35	34	31	32	33	31	32	33
90	28	30	34	31	31	30	32	30	32	32
75	25	25	31	25	27	26	28	29	29	29
50	21	22	22.5	23	24	23	25	24.5	24	23.5
25	16	18	19	15	20	20	18	19	18	18
10	14	15	14	14	17	16	16	17	16	27
5	12	14	13	12	16	12	16	15	16	13