INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6” x 9” black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700  800/521-0600
THE PERFORMANCE OF A MULTI-LINGUAL SOUTH AFRICAN SAMPLE ON
TWO MEASURES OF NONVERBAL ASSESSMENT

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree of Doctor
of Philosophy in the Graduate School of The Ohio State University

By
Andrew Martin, M. A., B.A.-Hons., B. A.

*****************************************
The Ohio State University
1996

Dissertation Committee:
A. Miranda, Ph. D.
G. Cartledge, Ph. D.
J. Collins, Ph. D.

Approved by

Adviser
College of Education
I would like to dedicate this document to my parents- Elizabeth and George Martin for being a source of strength. Your source of strength is the light at the end of the tunnel that keeps my hopes and dreams alive!
ACKNOWLEDGMENTS

Life is a tale of development. This document represents a milestone in my development. In the words of Andrae Crouch “I just want to take a little time and say, Thank You Lord for all HE has done for me”. Thank You for all the blessings You have bestowed on me. Without You none of this would have been possible.

A very special thanks goes to a remarkable woman, Dr. Antoinette Miranda for her unwavering patience and making me feel at home in her house. An educator in the true sense of the word, whose moral values are only surpassed by the loyalty and conviction shown to her advisees. I will miss her enthusiasm, warmth, smile, and appreciation for the field of school psychology- GOD BLESS you. Extra special “from the heart” thanks to Dr. Gwendolyn Cartledge for providing me with the strength at a time which I needed it the most. I thank her for being there and believing in me. Her patience and help through this learning experience was truly appreciated. To Dr. Collins, I thank him for his understanding and help. To Dr. Jim Statman thanks for taking time to personally make things happen. To Dr. Evelyn Luckey thanks for all her loving support. To Dr. Mary Dixon for her courage,
strength and integrity: thank you for your support through word and deed.

To Keith Wilson- thanks for providing me with those beautiful final color copy graphs.

To Margaret Price, her friendship, spirit and what she brought into my life... I have no words. Last, but not least, a special thanks to all those who have been there for me: Susan Boyer, Scott Gutentag, Randy & Lynette AU, Kathy Masica, Amanda Pickerill, Jan Powell, Laurice Joseph and Suzzane Ritter.
VITA

December 11, 1963................................................. Born - Cape Town, South Africa

1986........................................................................ B. A., Psychology
                          Western Cape University,
                          Cape Town, South Africa

1987........................................................................ Honors-B.A., Psychology
                          Western Cape University,
                          Cape Town, South Africa

1993.......................................................................... M. A., School Psychology
                          The Ohio State University

1988- 1991............................................................ Lecturer,
                          Western Cape University,
                          Cape Town, South Africa

1992-present.......................................................... Graduate Research
                          Associate,
                          The Ohio State University

1993- present........................................................ Research Assistant,
                          Otterbein College,
                          Columbus, Ohio

FIELDS OF STUDY

Major Field: Education

Studies in School Psychology and Psychology
Types of Schools in South Africa...............................................................35
  Private Schools..................................................................................35
  State Schools......................................................................................35

Procedures..............................................................................................37
  Naglieri Nonverbal Ability Test Answer Sheet....................................39
  Naglieri Nonverbal Ability Test Sample Items..................................40
  Stanford Achievement Test Answer Sheet........................................42
  Stanford Achievement Test Sample Items........................................43

Instrumentation........................................................................................44
  Naglieri Nonverbal Ability Test..........................................................44
  Stanford Achievement Test...............................................................47
  Stanford Achievement Reading Vocabulary subtest........................48
  Stanford Achievement Concepts of Number subtest........................48

Data Analysis............................................................................................49

IV. RESULTS................................................................................................51
  Summary of sample demographics...................................................51
  Summary of t-test analysis.................................................................52
  Summary of Descriptive Statistics of NNAT......................................54
  Summary of Descriptive Statistics of SAT.........................................54
  Summary of ANOVA of NNAT............................................................59
  Summary of ANOVA of Stanford Achievement Concepts of Number subtest..................................................59
  Summary of ANOVA of Stanford Achievement Reading Vocabulary subtest..................................................60
  Summary of Pearson’s Product Moment Correlations......................62
  Summary of ANOVA of NNAT by SES...............................................63
  Summary of ANOVA of Concepts of Number subtest by SES..........64
  Summary of ANOVA of Reading Vocabulary subtest by SES...........64
  Summary of Factorial Analysis of Variance of NNAT by Race and Socioeconomic status.........................................65
  Summary of Factorial Analysis of Stanford Achievement Concepts of Number subtest by Race and Socioeconomic status.....66
  Summary of Factorial Analysis of Stanford Achievement Reading Vocabulary subtest by Race and Socioeconomic status.....67

V. DISCUSSION.............................................................................................69
  Summary of Study...............................................................................70
    Within group differences on NNAT by gender..............................70
    Between group differences on NNAT by race...............................70
    Within group differences on SAT by gender...............................73
    Between group differences on SAT by race.................................73
    Relationships between NNAT and SAT by race and total sample....75
Limitations of the Study and suggestions for future research........ 76
Conclusion..................................................................................................... 78

LIST OF REFERENCES...................................................................................... 83
Appendices............................................................................................................. 92
## LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Group, grade, age, and gender characteristics of sample</td>
<td>37</td>
</tr>
<tr>
<td>2. Sample Demographics</td>
<td>52</td>
</tr>
<tr>
<td>3. Summary of T-tests analysis between groups by gender of NNAT and selected Stanford Achievement subtests</td>
<td>53</td>
</tr>
<tr>
<td>4. Means and Standard Deviations of NNAT by group and grade</td>
<td>54</td>
</tr>
<tr>
<td>5. Means and Standard Deviations of selected Stanford Achievement subtests by group and grade</td>
<td>56</td>
</tr>
<tr>
<td>6. Summary of Analysis of Variance of NNAT by race and grade</td>
<td>59</td>
</tr>
<tr>
<td>7. Summary of Analysis of Variance of Stanford Achievement Concepts of Number subtest by race and grade</td>
<td>60</td>
</tr>
<tr>
<td>8. Summary of Analysis of Variance of Stanford Achievement Reading Vocabulary subtest by race and grade</td>
<td>61</td>
</tr>
<tr>
<td>9. Obtained Pearson r Correlations between NNAT and selected Stanford Achievement subtests by group (each group ( N = 100 )) and Total scores for Total Sample (( N = 400 ))</td>
<td>62</td>
</tr>
<tr>
<td>10. Summary of Analysis of Variance of NNAT by SES</td>
<td>63</td>
</tr>
<tr>
<td>11. Summary of Analysis of Variance of Stanford Achievement Concepts of Number by SES</td>
<td>64</td>
</tr>
</tbody>
</table>
12. Summary of Analysis of Variance of Stanford Achievement Reading Vocabulary subtest by SES............................................. 65

13. Summary of Factorial Analysis of Variance of NNAT by race and socioeconomic status....................................................... 66

14. Summary of Factorial Analysis of Variance of Stanford Achievement Concepts of Number subtest by race and socioeconomic status................................................................. 67

15 Summary of Factorial Analysis of Variance of Stanford Achievement Concepts of Reading Vocabulary subtest by race and socioeconomic status....................................................... 68
LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Naglieri Nonverbal Ability Test Answer Sheet</td>
<td>39</td>
</tr>
<tr>
<td>2. Naglieri Nonverbal Ability Test Sample Items</td>
<td>40</td>
</tr>
<tr>
<td>3. Stanford Achievement Test Answer Sheet</td>
<td>42</td>
</tr>
<tr>
<td>4. NNAT Mean Raw Scores by Grade and Race</td>
<td>55</td>
</tr>
<tr>
<td>5. Stanford Achievement Reading Vocabulary subtest Mean Raw Scores by Grade and Race</td>
<td>57</td>
</tr>
<tr>
<td>6. Stanford Achievement Concepts of Number subtest Mean Raw Scores by Grade and Race</td>
<td>58</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

By the turn of the 19th century we were introduced to the term IQ by a French psychologist, Binet (1916) who developed the first IQ scale. Binet's goal was to use the instrument to demonstrate differences between individuals only if they had more or less the same opportunities (Maurice, 1989). Unfortunately, Binet's work created a window of opportunity for a certain stratum of psychologists in the United States. It was also during this time that we witnessed the changing nature of capitalism, thus producing a certain stratum who would in essence represent the historical continuity of social life (i.e. Cattell; Terman; Thorndike; Yerkes; etc, etc). The custom was to alter social contradictions without threatening existing relations. This became a supreme method of sustaining capitalism. To ensure the continued existence of capitalism, a layer of stratum was pivotal to this process. Henceforth, attempted theoretical work related to intelligence tests started to make their appearance to substantiate their viewpoints. Two schools of thought became known. First, the view that intelligence is best described as a single general
factor (Spearman), and secondly, that intelligence is multi-factored (Guilford).

Brody (1992) summarized Spearman’s theoretical frame of mind and perception of intellectual functioning very well. Spearman asserted intelligence could be divided into two components, general component (g) and specific component (s). The g component is determined by that which the measure has in common with other measures of intellectual functions. The s component is specific to each measure. Thus, implying that the correlation between two measures of intelligence will be determined by the ratio of g-to-s in each of the measures. This theory became known as the two-factor theory of intelligence (Brody, 1992).

Guilford (1964; 1967; 1985; cf. Brody, 1992) based his work on a three-dimensional model of intellectual tasks. These mental tasks according to Guilford differed with regard to mental operations, contents and products. He asserted that there were five operations: cognition (knowing), memory, divergent production (generation of alternatives), convergent production (generation of logic-tight conclusions) and evaluation. These operations can be applied to four types of contents, figural, symbolic, semantic and behavioral. Guilford (1977) subsequently modified his work and came to the conclusion that the application of a particular operation to a particular content can be expressed in terms of one of six products. These are units, classes, relations, systems, transformations, and implications. This theory implies that there are 150 different types of intellectual tasks.
Since the inception of the aforementioned schools of thought, many tests have been developed and standardized (e.g. Wechsler Intelligence Scale for Children; Wechsler, 1974). However, these instruments have been criticized for a number of reasons. Matthews (1980) summarizes the criticism of intelligence tests beautifully. *First*, IQ tests are validated on achievements, which are socially predetermined. Hence, IQ tests begin with the identification in advance of groups that are accepted as intelligent and unintelligent. Therefore, these then become the groups against which the tests are validated. *Secondly*, the raw material of the “science” is an abstraction from the real world, it socially produced, theoretical object- the IQ score. This means it has no concrete properties itself but is surmized from the behavioral tasks. *Thirdly*, judgements of intelligence are comparative and relative rather than absolute. The use of these tests have survived in one human environment, namely the “white, middle-class” environment (Kaplan & Saccuzzo, 1989). It became clear that IQ tests were and continue to be dismally ineffective, since it focuses on differences among individuals (Valdes & Figueroa, 1994). There have been several court cases that have challenged the use of IQ tests with culturally diverse populations. For example, Diana v California State Board of Education (Valdes & Figueroa, 1994). This case had two implications. *First*, children had to be tested in their primary language as well as in English. *Secondly*, nonverbal IQ tests could be utilized as a substitute for full-scale IQs in assessing Hispanic pupils with limited English.
Another litigation was Larry P v Riles (Valdes & Figueroa, 1994). In this case the Judge proposed that the use of IQ measures for assessing black children considered eligible for EMR classes be banned. The Judge made the following recommendations: *First*, to use IQ measures the California State Board of Education needs to get approval from the court. *Secondly*, the court will ensure that the measures which are being used on black children were not racially or culturally discriminatory. *Finally*, all black children in EMR classes had to be retested. During this decade of court cases as well as other similar litigated courts cases, we witnessed the emergence of Public Law 94-142.

As a result of these controversies we have seen the use of “nonverbal tests”, which are more commonly referred to as “culture-fair” tests (e.g. Raven’s Progressive Matrices, 1947; Wechsler Performance Scale; 1974; Matrix Analogies Test; 1985; etc.). According to these proponents nonverbal tests have several uses. *First*, they can be used to evaluate intellectual levels of children from diverse educational and linguistic backgrounds. *Second*, they help identify children who may experience discrepant performance between their nonverbal reasoning ability and achievement. *Third*, they can be used for the nonverbal assessment of ability in exceptional populations such as deaf children with communicative disorders and the mentally retarded. *Fourth*, they assist in the assessment of children with motor problems since the instrument does not require the manipulation of concrete materials.
Finally, they may be used as a screening instrument in the early identification of students who, after complete psychoeducational evaluation, may qualify for special educational services (Bardos, 1990). Proponents have argued that nonverbal tests are easily transportable across cultures (Jensen, 1980). While there is a recognition that the use of nonverbal tests is a valid attempt at addressing the problem of assessment with culturally diverse populations we must not be misled by the thought that it is a full proof method. “Culture-fair tests” have been associated with a number of problems, especially with populations from diverse cultural backgrounds. Nonverbal measures of intelligence have always shown lower predictive validities, and data suggest that they may even further compromise predictive validity by being hypersensitive to language (Figueroa, 1990). There appears to be a failure to take into account the different types of minorities and the differences in the relationship between their cultures and the mainstream American culture (Ogbu, 1992). Lonner (1990) has asked if a concept developed in one culture should be transferred to another culture as if the concept will not change in the process.

**Summary**

Nonbiased assessment is an extremely complex issue. Anastasi (1988) indicated that the search for a culture-free test is an illusion and culture cannot be ignored. Until we have have produced assessment measures that are fair
and common to all cultures, we need to continue with the close scrutiny of assessment instruments (Kaplan & Saccuzzo, 1989).

**Statement of the Problem**

The problem faced by those who wish to use such nonverbal measures in South Africa demand great resourcefulness. The major purpose of the present investigation is to examine the utility of a “culture-reduced” test as a screening instrument for children in South Africa who are in educational settings. Certain researchers claim that these measures of general intelligence are culturally reduced, and that they are more appropriate for diverse cultural groups (Jensen, 1980). It is the aim of this investigation, therefore, to determine how different population groups perform on these nonverbal measures of intelligence, the Naglieri Nonverbal Ability Test* and the Stanford Achievement Test (SAT; 1988).

Specifically, this investigation will examine the predictive validity of the NNAT using a group administered achievement test (SAT) as a criterion.

The following Research Question will be investigated:

1. Performance of Black, White, Coloured and Indian students on the NNAT:
   a. Is there a significant difference between groups, by gender within each group?

*Note: (*) An asterisk indicates an unpublished experimental form.*
b. Is there a significant difference between groups, by race?

2. Performance of Black, White, Coloured and Indian students on the SAT:
   a. Is there a significant difference between groups, by gender within each group?
   b. Is there a significant difference between groups, by race?

3. Overall relationship between the NNAT and SAT:
   a. What are the relationships between the NNAT and SAT for the entire sample and by race?

Definitions of Terms

Following are definitions of ethnic groups which will be used in this research study.

**White**

This community’s history can be traced back to the first Dutch settlers at the Cape of Good Hope in 1652. Over a period of time, this was followed by a huge influx of immigrants from Germany, France, Portugal, Greece and Italy. These immigrants later formed a distinct cultural group in South Africa which became known as the Afrikaners. Today, Afrikaners comprise of 60% of South Africa’s white community with Afrikaans as their native language. Afrikaans is unique to South Africa, in the sense that it originated from South Africa. Afrikaans is a combination of Flemish and Dutch.

7
The other 40% are English speaking, who are descendants of Britain. Both Afrikaans and English languages been promulgated by the government.

Coloured

The so-called “coloured” community is a product of interbreed between the earlier tribes of the Cape, white settlers and a substantial amount of slaves who had been imported from the East and West Africa. The amalgamation of these different cultural groups produced children and they became known as the “coloured” people. The “coloured” community has two distinct subgroups, the Cape Malays and the Griquas. The Cape Malay have a stern Muslim tradition and reside in Cape Town. The Griqua community can be found in the North-Eastern part of the Cape Province. Greater Cape Town is being absorbed by the “coloured” community with Afrikaans as their native language. With the exception of the Muslim community, most of the “coloured” community are Christians and their lifestyle is more geared toward a western trend.

Indians

The majority of the Indian community originates from India. These people were brought to South Africa as early as 1860 to work as laborers on the sugar plantations in Natal. Natal is the North Eastern Province of South Africa. Currently it is estimated that 80% of the Indian community still resides in Natal. They are fluent in English and have a very high literacy rate and are well-educated. Their lifestyle is a combination of western and oriental customs.
Blacks

The black community is not a single homogeneous group. They are comprised of eight major distinctive ethnic groups which are the Zulus, Xhosas, North Sotho, Southern Sothos, Tswana, Venda, Tsonga and Swazis.

Significance of the Study

The significance of the present study is that it is the first time that an American normed test has been used on a South African population to possibly determine the feasibility of adapting such a measure to South African cultural groups. The test's author describes it as a nonverbal test, however, current measures of nonverbal tests do not adequately cater to groups from culturally diverse backgrounds.
CHAPTER II

REVIEW OF THE LITERATURE

Binet (1905), a French psychologist's greatest contribution to the field of psychology was the development of the first psychological scales to measure intelligence (Hothersall, 1990). Binet was aware of the limitation in the application of his method and pointed out that his tests could safely be used to determine differences between various individuals or groups if they had more or less the same opportunities (Maurice, 1989). Spearman criticized Binet's research by asserting he failed to utilize laboratory measures in his research. However, he recognized that Binet's tests could provide an adequate measure of intelligence (Spearman; 1927; 1930; cf. Brody, 1992). Spearman agreed with Binet's tests on the notion that a single score must be interpreted as g (Brody, 1992). Spearman believed that his theory provided the foundation for the measurement of intelligence, introduced by Binet. Thus Binet provided a test and Spearman provided a theory (Brody, 1992).
THEORIES OF INTELLIGENCE

Spearman's General Factor Theory

Galton, Spearman and Guilford have in one way or another contributed to intelligence. It was Galton who initially suggested exploring the relationship between measures of sensory discrimination and how it relates to outcome of individual differences of intelligence. Galton concluded in his research that sensory functioning correlates significantly with measures of intellectual functioning (Brody, 1992). Galton's research drew attention from many researchers, especially Spearman. The latter went on to question Galton's results (Brody, 1992).

Spearman's critical points of Galton are summarized by Brody (1992) as follows: *First*, overreliance on subjectivity-noting that formulas for the calculation of the coefficient of correlations existed and that this statistical measure could be used to obtain the degree of linear relationship between the two sets of measures. *Secondly*, no evidence exist's of statistical variance in his results. *Thirdly*, the research study was very poorly managed. The probability existed that the results could have been influenced by several variables. *Fourth*, the unreliability of the measurement could not be amended (Brody, 1992). Spearman's theoretical assumption was, if sensory discrimination influenced intelligence, there should be a significant relationship (Brody, 1992).
This relationship accordingly contained two components, a general component (also known as g) and a specific component (also known as s). Spearman referred to this theory as the two-factor theory (Brody, 1992). Inevitably, Spearman’s view of intellectual functioning had an impact on the research on the measurement of intelligence (Brody, 1992). Spearman was convinced that there appeared a frequent significant correlation among different measures of intelligence (1904, cf. Brody, 1992).

Spearman’s theoretical frame of mind and his perception of intellectual functioning is summarized by Brody (1992) as follows: First, the intercorrelation of test scores should display a positive manifold, which in essence means that all of the correlations should be positive. This follows directly from the assumption that all measures of intelligence have a g component. Secondly, a positive hierarchy should be visible when you view the g-to-s ratio. Since Spearman assumed that the correlation between any two measures of the common intellectual function is determined by the amount of g they share, it follows that the measure with the highest average correlation with all other measures will have the highest ratio of g-to-s. A rank ordering of the measures may be obtained from the average correlation of each measure with all other measures. Thirdly, the matrix needs to fulfill an important law, Spearman referred to it as the law of tetrad differences (Spearman & Holzinger, 1924, cf. Brody, 1992).
The law may be explained as follows:

\[ \Gamma_{12} \times \Gamma_{34} = \Gamma_{13} \times \Gamma_{24} \]

Drawing on Spearman's literature it was clear that his theme of intelligence was based on cognitive principles. They are called "eduction of relations and correlates" which Spearman clearly defined as the following:

The eduction of relations...when a person has in mind any two or more ideas (using this word to embrace any items of mental content...) he has more or less power to bring to mind any relations that essentially hold between them. It is instanced whenever a person becomes aware, say that, that beer tastes something like weak quinine...or that the proposition "A is B" proves the proposition "Some A is B"...

The eduction of correlates...when a person has in mind any idea together with a relation, he has more or less power to bring up into mind the correlative with a relation, he has more or less power to bring up into the correlative idea.

For example, let anyone hear a musical note and try to imagine the note to fifth higher.. (pp. 165-166, Spearman, 1927 cf. Brody, 1992).

The aforementioned example constitutes Spearman's theory of general intelligence. Another theory of intelligence (multi-factored) was proposed by Guilford.

**Guilford's Multi-Factored Theory**

Guilford (1964; 1967; 1985; cf. 1992) based his theory on a three-dimensional model of intelligence tasks. These tasks differ from each other with regard to mental operations, contents and products. Guilford also asserted that there were five operations- cognition, memory, divergent production, convergent production and evaluation. An application of a specific operation to a specific content can be expressed in terms of six products. These products
are units, classes, relations, systems, transformations, and implications.

Guilford’s theory implies that there are 150 different types of intellectual tasks. Thus, a model of 150 independent ability factors.

Guilford used factor analysis to prove his theoretical point of view. These factor analyses provided him with the evidence for the independence of the abilities postulated in the model. Guilford (1964) rejected factor g on the basis of his finding, which were that 17% or 48, of 140 correlations among the tests of abilities in his study fell within the -.10 to +.10 range. However, it should be noted that Guilford’s results does not support his theory (Brody, 1992). The latter pointed out that Guilford’s reported 83% of his correlations among ability measures exceeded .10. If the assumed theory were to be correct the ability measures would usually turn out to be zero.

In light of this, Brody (1992) warns that Guilford’s findings should be interpreted with caution, for three reasons. First, some of the tests included in his matrices had low reliability. Thus, unreliable tests may not correlate highly with themselves, nor are they to correlate highly with other tests. Secondly, Guilford frequently used Air Force officer trainees as subjects. They were selected for intelligence purpose. Thus, his sample was restricted. Thirdly, some of the abilities in his model may not be related to general intelligence. Especially tests with behavioral content may be inadequate measures of social skills. It should be noted that Guilford’s theory has no empirical support.
Thus, the postulated factor structure appears to be a myth (Brody, 1992). With the inception of these theories of intelligence we witnessed the development of intelligence tests.

**Intelligence Tests: Uses and Implications**

Interest in intelligence testing grew out of the recognition of individual difference in people and the need for academic screening. One of the most popular intelligence tests is the Wechsler Intelligence Scale for Children (WISC; developed by David Wechsler, 1949). The upshot regarding the development of these intelligence tests is to measure the amount of learning skills an individual has acquired in the curriculum of the given culture. When an individual takes an intelligence test, he or she is measured relative to the amount of acquired knowledge and learning skills in the school and in comparison to others of the same age. Thus, the assumption is the person taking the test comes from the same “classroom” had been exposed to the given curriculum, has acquired the same general motivation towards solving the test items, and has had about the same types of experiences with test taking. Taking these factors into account, the IQ score is supposed to inform the examiner of the individual’s learning ability and “intelligence”.

One of the most enthusiastic proponents of IQ, Arthur Jensen (1977), made the observation that there appears to be an empirical relationship between intelligence and learning. Intelligence is crucial for learning that is
intentional, requires use of hierarchically arranged knowledge, requires comprehension, relies on transference, is insightful, is complex, must be done within a specified and limited time frame, represents a new problem, and is keyed to chronological age. The crucial issue in these relationships is the applicability of the assumptions about an individual's background and his experience with the classroom. There has always been a concern about the differences of tests scores among individuals from different cultural backgrounds. Intelligence tests have been used in cultures other than those on which they were standardized for as long as the history of testing itself. Furthermore, there has been widespread concern about the applicability of tests across culturally diverse groups for years (Ahia, 1984; Groth-Marnat, 1990; Kaplan & Saccuzzo, 1989).

Figueroa (1990) identified four parameters regarding the use of IQ tests on culturally diverse populations. First, the majority of language minority groups show the same, low verbal IQ, high nonverbal IQ profile. Secondly, those who perform poorly on verbal is regarded as having a language handicap. Thirdly, psychometric properties are used as indices of test appropriateness for use with diverse populations. Fourth, most studies ignored socioeconomic status and the effects of inadequate schooling. The literature on intelligence tests have not changed not that much with regard to the treatment of culturally diverse populations (Jensen, 1980, 1984; Kamphaus & Reynolds, 1987; Sattler, 1988, 1992). Interestingly, the notion of
bias in measures of intelligence have changed little as well. However, promising trends are evidence in studies (Clark, 1987; Miller-Jones, 1989). We witnessed during the early 1970s a number of law related court cases on the misuse of assessment with culturally and linguistically different children. Particular attention will be paid to two cases which had far reaching implications for the use of assessment on culturally diverse populations. They are Diana v State Board of Education and Larry P v Riles. The following description of the latter two court cases were taken from a book: Bilingualism and Testing: A Special Case of Bias, by G. Valdes and R. Figueroa (1994).

**Diana v California State Board of Education (1970)**

In 1969 the parents of nine Mexican American children who had been placed in educable mentally retarded (EMR) classes in Monterey County in California filed a law suit against the State Board of Education. These parents felt that their children were short changed by placing them in EMR classes. To add fuel to the fire, the EMR final test score results proved to be invalid. The children spoke predominantly Spanish, yet they were given an IQ test in English.

Diana’s law suit never went to court, mainly because the nine plaintiffs was a simple extrapolation of what was happening throughout the State of California. Most of these classes were heavily overrepresented by Hispanic pupils. The out-of-court settlement was arrived at before Judge Peckham.
This case established a unique set of rules regarding the use of IQ test on limited English proficient (LEP) pupils. For example, testing was to be done in the children's primary language and in English as well, nonverbal IQ tests could be used on those children who had been previously diagnosed as EMR. Chicano children had to be retested and special attention had be given to their nonverbal scores. An IQ instrument that is appropriate for Mexican-American children had to be developed, and the representation of EMR Hispanic students in each school district throughout the State of California had to be monitored to make sure there were no ethnic disparities of Chicano pupils in EMR classes.

This agreement was problematic for two reasons. First, the monitoring agreement was seen as a quota. Secondly, the manner in which a district was judged regarding overrepresentation of EMR children became quarrelsome. In June 1973, the court set fourth a stipulation wherein "E Formula"* was proposed as a vehicle for determining when a district was overrepresented. At a meeting in August of 1973, the State Department of Education rejected the plaintiffs' "E formula" for determining significant variance. The upshot for the department's rejection to the plaintiffs' formula was the Larry P case. Larry P counsel's objective was to obtain a similar settlement to that of the Diana case.

\[ E = A + \sqrt{\frac{A(N - A)}{N}} \]

where \( A \) represents the percentage of Spanish surnamed pupils in a district, \( N \) represents the total number of all pupils in the EMR classes of a district, and \( E \) represents the maximum percentage of Spanish surnamed pupils permitted in EMR classes.
The implication would be the removal of misplaced black children from EMR classes. On May 24, 1974, Judge Peckham settled the dispute. In a memorandum and Order, the Judge stipulated that the 1973 "E Formula" was to be used. The following sections 5, 6, and 7 were stipulated in the 1973 memorandum. They were as follow: Section 5: If a significant variance persists after a period of three years, the State Department will conduct a thorough investigation, which may include reevaluation of pupils if necessary. Section 6: If any inequalities existed after a period of three, the State Department may investigate the district by sending a monitor or a review team to ensure that their methods are appropriate. Section 7: If during any period of time a variance exists in a district, which did not exist at the time when the stipulation was reached, the particular district will immediately be contacted. Another case of similar nature was the Larry P case.

Larry P case v Riles (1970)

The Larry P case went to Judge Peckham's court in 1970. The facts of the case were very much similar to those of Diana. A small group of black children who had been in EMR classes brought a law suit against the San Francisco Unified School District and the State of California because they were misplaced on the basis of biased, invalid IQ tests. In 1972, the judge temporarily suspended the use of IQ test measures which did not take into account black children's background and cultural experiences.
In 1979, the judge banned the use of IQ tests with black children who had been referred for EMR placement. The 15 IQ-point discrepancy between black and white children were not explained in terms of genetic differences by any of the witnesses for the plaintiffs, and only two witnesses for the defense. Neither environment nor genetics served to explain this difference, thus the notion of test bias became critical. David Wechsler and Wilson Riles, has led the judge to conclude that “there was general agreement by all sides on the inevitable effect of cultural differences on IQ scores” (p. 48). The judge suspended the use of IQ measures for assessing black children considered eligible for EMR classes. The following proposal was made by the judge: To use IQ measures the California Board of Education needs to get approval from the court. Approval will only be granted if the following are met. First, that the test were not racially or culturally discriminatory. Secondly, that it would not lead to overrepresentation of black children. Thirdly, that these placements were validated. Finally, data needs to support the aforementioned, and public hearings had to be called with regard to the state’s intention of using these tests. The state had the right to monitor district overrepresentation using a procedure which seems essentially identical to the Diana case “E formula”. All black children in EMR classes had to be reevaluated without using unapproved IQ measures. This reevaluation should include the following. First, diagnostic tests designed specifically to reveal learning needs and should prescribe
specific teaching approaches. *Secondly,* the reevaluation should include adaptive behavior observation. *Finally,* the reevaluation should also include developmental and health histories.

Court cases such as Diana *v* State Board of Education and Larry P *v* Riles, became the sole direct reference to meeting the assessment needs of minorities- Public Law 94-142. In light of Public Law 94-142 (Federal Register, 1977), the legislation mandating psychoeducational assessment and services, we witnessed the validation of nonverbal intelligence measures for specific use with minority children (Valencia, 1982).

**Nonverbal Intelligence Tests**

Measures of nonverbal intelligence should possess three essential characteristics. *First,* the instrument must limit the degree to which the examinee is required to recall or demonstrate another base of knowledge (Lohman, 1979). This restriction is accomplished by either administering the instrument through nonverbal means (e.g., in a pantomimed fashion) and/or by ensuring that the test is a homogeneous measure of nonverbal abilities. Solution to nonverbal tasks should not depend upon the individual’s previous experiences. For example when required to solve problems that are figural in nature, the examinee should not have to recall the vocabulary of a specific subject matter. Instead the subject should apply reasoning strategies that are
unique to the problem presented by the figures. Secondly, the test must require
the examinee to use strategies that involve complex reasoning. Individual
differences in reasoning are not observed until the items become challenging to
the examinee (Bethel-Fox, Lohman, & Snow, 1984; Sternberg & Gardner, 1983).
Snow, Kyllonen and Marshalek (1984) used a scaling solution to examine the
relationship among ability tests and examined various content (figural, verbal
and numerical) as well as the complexity of the content. Tests that included
analogies, concept formation, and block designs were identified as being most
complex. Complexity also implies novelty, and both novelty and complexity
make a nonverbal test more process-intensive (Lohman, 1988). The examinee
must focus on how to solve the problem rather than the content of the problem
itself. Thirdly, assessment of nonverbal intelligence should require flexibility in
the examinee’s application of reasoning strategies. Individuals’ may employ
various strategies to solve the same problem and may change strategies as
problems increase in difficulty (Snow & Lohman, 1988). The examinee employs
various strategies to improve his performance on nonverbal measures of
intelligence.

Literature points out that the theory of nonverbal intelligence can be
found in Spearman and Cattell’s work. In 1927 Spearman hypothesized that a
general ability (g) is present in all ability tests. This general ability factor is
assessed in nonverbal tests by abstract or complex reasoning tasks. Cattell
asserted that Spearman’s (g) can be divided into two component, fluid and
crystallized intelligence. This two-tailed theoretical dimension of intelligence displayed the ability for abstract, conceptual formulation and perceptual and eductive relations (Gustafsson, 1994). Cattell identified abstract reasoning as being related to fluid intelligence or cognitive problems. Fluid abilities represent the cognitive skills that are necessary to solve novel problems involving stimuli. General experiences are considered more important in solving these types of problems than schooling. In short, "this factor involves the intervention of new cognitive strategies or the flexible reassembly to deal with novel situations" (Thorndike et al., p. 4, 1986). Crystallized abilities are influenced by schooling as well as general life experiences that are acquired outside of school. Horn (1976) stated that fluid intelligence appeared to be influenced by educational experience and crystallized intelligence is influenced by cultural experiences. Thus nonverbal intelligence tests is a measure of fluid and crystallized intelligence.

**Uses of Nonverbal Tests**

Nonverbal tests offer an alternative to those who have been difficult to assess, for example, individuals who have speech or language related disabilities, hearing impaired, etc. In addition individuals who do not have literacy skills or who have motoric impairments may also benefit from nonverbal assessment because the examinee is not required to perform paper/pencil tasks. Nonverbal measures of intelligence is also helpful with
those individuals who are illiterate, children who are bilingual or who have limited language skills. The use of nonverbal measures intelligence is to assess the intelligence of LEP children is a recommended practice (Figueroa, 1990). Nonverbal measures of intelligence have been considered useful in assessing minority students, since they are viewed as nonbiased and less culturally loaded than language tests. However, this assertion is not supported by the literature for black children (Anastasi, 1982). Following are a brief description of selected nonverbal measures of intelligence (Raven's Progressive Matrices; Leiter International Performance Scale; Hiskey-Nebraska Test of Learning Aptitude; Test of Nonverbal Intelligence; Kaufman Assessment Battery for Children-Nonverbal Scale; Wechsler Intelligence Scale for Children-Performance Scale IQ and Naglieri Nonverbal Ability Test).

The Raven's Progressive matrices is the most widely used measure of nonverbal assessment. The progressive matrix format uses abstract figural diagrams. These are typically arranged in a 2 by 2, 2 by 3 or 3 by 3 configuration. In order to solve the problem, the examinee must evaluate the information provided, gain an understanding of the changes that are printed in the matrix, and determine which of the several options correct completes the matrix. There are two Raven's Progressive Matrices which are used to measure nonverbal intelligence. They are Standard Progressive Matrices (Raven, 1947a), Coloured Progressive Matrices (Raven, 1947b). The most frequently used form is the Standard Progressive Matrices.
The Leiter International Performance Scale (LIPS; Leiter, 1979) is a 54-item scale of general intelligence, and was standardized in 1948 on children between the ages of 2 and 18 years (Naglieri & Prewett, 1989; Sattler, 1992). The instrument consists of several categories, which include perception, symbolic transformation, quantitative discriminations, spatial imagery, genus matching, progression discriminations and immediate recall. These are organized into two major groups, perceptual and conceptual items (Johnson, 1982). The LIPS is composed of wooden blocks with the subject’s objective being to match these blocks with each other (Naglieri & Prewett, 1989).

There are four items at each mental age from 2 years old to 10 years old, two items per year from 11 year olds to 16 year olds and six items in the adult range (Johnson, 1982). It has administration time of approximately 45 minutes (Sattler, 1988). The test has been frequently utilized to assess intellectual functioning of bilingual (Gerken, 1978), hearing-impaired (Spellacy & Black, 1972), aphasic (Black, 1973) and brain-injured children (Bensberg & Sloan, 1951).

The Hiskey-Nebraska Test of Learning Aptitude (H-NTLA; Hiskey, 1966) is a performance scale (Sattler, 1988) primarily developed to assess the intellectual functioning of hearing-impaired children and adolescents between the ages of 3 and 17 years old (Salvia & Ysseldyke, 1981). There are 12 individual subtests which measures several aspects of intelligence such as visual memory, visual discrimination, visual organization, and visual association (Watson, 1983). An important characteristic of the scale is the
parallel administration of the subtests (Sattler, 1988). Administration occurs through pantomimed directions (Watson, 1983) and a nonverbal response is required from the testee (Sattler, 1988). The approximate amount of time to administer the instrument is 50 minutes (Hiskey, 1966).

The Test of Nonverbal Intelligence (TONI; Brown, Sherbernou & Dollar, 1982) is a 50-item measure of nonverbal intelligence with two equivalent forms, A & B, primarily designed as a language-free measure of cognitive ability and a global measure of intelligence (D'Amato & Lassiter, 1994). Since it is fundamentally based on problem solving, the TONI "... lends itself readily to the abstract content and nonverbal testing format" (Brown et al., p.12, 1982). The TONI should be individually administered to subjects between the ages 5 years and 85 years old with an administration time of 15 minutes (Brown et al., 1982).

The Kaufman Assessment Battery for Children (K-ABC; Kaufman & Kaufman, 1983) is a fairly recently developed, individually administered measure of intellectual functioning and achievement measure for children between the ages 4 years 6 months and 12 years 6 months. The instrument is composed of 16 subtests. These subtests are grouped into three global areas (Sequential Processing Scale, Simultaneous Processing and Achievement Scale). The instrument contains a Nonverbal scale for those who display a communication disorder (Kennedy & Hiltonsmith, 1988). The K-ABC Nonverbal scale consists of three to five selected subtests, depending on the
subject's age (Naglieri & Prewett, 1989). Following is a brief description of the five subtests (Porter & Kirby, 1986)  

*Hand Movements:* Measures the individual's ability to replicate the exact sequence of taps on the table with the fist, palm, or side of the hand after the examiner's example.  

*Triangles:* Measures the person's ability to assemble several identical triangles to form different designs from pictures.  

*Matrix Analogies:* Measures a individual's ability to select a design that best completes a visual analogy.  

*Spatial Memory:* Measures memory.  

*Photo Series:* Measures the individual's ability to place photos in proper sequence.  

The total Nonverbal K-ABC score is reported as a standard score with a mean of 10 and a standard deviation of 3. Each subtest yields standard score of 10 (Porter & Kirby, 1986). Since only a single sequential processing subtest is included within the scale at each age level, the Nonverbal scale yield a single global score (Kennedy & Hiltonsmith, 1988). The sequential processing component for the school-age children does not exist (Naglieri & Prewett, 1989). The K-ABC NVs offers a valuable tool in the cognitive assessment of speech-and-language-impaired children (Telzrow, 1984).

The Wechsler Intelligence Scale for Children Revised (WISC-R; Wechsler, 1974) contains 12 subtests of which six subtests are on the Verbal Scale and six are on the Performance Scale. The Wechsler Intelligence Scale Third Edition (WISC-III; Wechsler, 1991) has two supplementary subtests, Digit Span (Verbal Scale) and Mazes (Performance Scale). The instrument is an individually administered measure of intellectual functioning between the ages 6 years
through 16 years, 11 months. The Performance Scale IQ is calculated on the basis of five Performance subtests (Picture Completion, Picture Arrangement, Block Design, Object Assembly and Coding).

The Naglieri Nonverbal Ability Test (NNAT) is comprised of 38 multiple choice (5 options) items and is intended to be administered individually or by group. The test is intended to be a general nonverbal measure of intelligence like other tests that utilize the standard progressive matrix format (Naglieri, 1985). The NNAT is composed of items of the types found in the Ravens Progressive Matrices (Carlson and Wiedl, 1976). The test includes items in groups called Pattern Completion (requires the individual to choose the options which accurately completes the pattern), Reasoning by Analogy (requires the examinee to see how the change or changes in one figure relates to the analogous change or changes in another), Serial Reasoning (requires the individual to discover the order in which items appear throughout the matrix) and Spatial Visualization (requires the individual to imagine how a figure would look like when two or more designs are combined). The NNAT can be administered on children between ages 5 to 17 years old. The test is useful with linguistically and cultural diverse populations, since the administration requires minimal verbal directions.
Although nonverbal measures have been widely used among culturally diverse populations, for example, Argulewicz, Anderson and Bingenheimer (1983) with Peabody Picture Vocabulary Test and the Raven Progressive Matrices; Sandoval (1979), Elliott et al., (1985), Elliot and Boeve (1987) with the WISC-R; Valencia (1985a, 1985b) with Kaufman Assessment Battery for Children; Powers, Barkan and Jones (1986) with the Raven Progressive Matrices; Elliot, Piersel, Witt, Argulewicz, Gutkin, & Galvin (1985) with the WISC-R, the successful outcome of these measures has been limited. Adding an additional dimension such as cross-cultural makes the problem even more difficult. The problem becomes especially more difficult when we want to employ measures which were standardized on a first world society and transfer them into a third-world society. Methodological issues in cross-cultural research are varied and complex.

Cross-Cultural Use of Nonverbal Tests

The reasoning behind comparative research is to assess different people in a comparable manner. However, when the assessment involves cultural groups with diverse cultural backgrounds, the attempt to find a common ground becomes more difficult.
The comparison of one or more cultural groups in an attempt to determine their level of cognitive ability on any given nonverbal intelligence task needs to be cognizant of the perceived differences which are not simply due to "familiarity" with the materials (Glick, 1984). Various authorities concluded from their reviews that comparisons between those who are familiar versus those who are unfamiliar with materials may be due to classification abilities with those subjects who were better familiar with the materials (Okonji, 1971; Greenfield, Reich, and Olver, 1966). Research also points out the increase probability of finding differences in performances based upon the schooling factor. These differences may, in fact, be related more to the "familiarity" of the task to the schooled group than any cognitive impact of schooling (Glick, 1984). Communication problems exist when we introduce a task, establishing clear communication of instructions, establishing clear translation of responses which may establish unintended attitudes toward the task. Emphasized in the literature were experimental test items or instructions (Werner & Campbell, 1970; Sechrest, Fay, and Zaidi, 1972). Tasks requiring nonverbal responses usually use a verbal introduction. Therefore, problems of linguistic and developmental comparability of task instructions must be dealt with (Glick, 1984). Cross-cultural research also involves motivational problems. The problem could influence people's performance on cognitive tasks (Glick, 1984). Biesheuval (1949) indicated that the establishment of motivational equivalence is a difficult one for cross-cultural research.
When cultures differ, for example, in such factors as competitiveness, it makes it difficult to assure equal motivation in different samples (Madsen, 1971). Investigators have also noted, moreover, that the quality of motivation may vary widely between cultures. For example, schools in developing areas (where tribunal linguistic differences exist) emphasized rote memorization and strict obedience to the teacher (Gay & Cole, 1967).

Observations were also made by Greenfield and Bruner (1969) who offered an explanation in terms of cultural value system (collective versus individualistic). Thus, the assumption may not be made that figural and perceptual thinking is a-cultural. Perceptual, learning and reasoning processes are best understood within a culture's distinct frame of reference. Furthermore, these dimensions are being influenced by the interactions which inturn determine the skills that is being learned and these learning skills are controlled by the cultural group's value and social system. Culture equals thinking (Vygotsky, 1987) and is manifested in social praxis.

Standardized measures of nonverbal intelligence rely upon two-dimension representation of three-dimension objects (perceptual tasks involves two-dimensional information Glick; 1984). The use of either representational or abstract two-dimensional presentation with non-western populations are problematic (Biesheuval, 1949; cf. Price-Williams, 1970). In a study conducted by Mundy-Castle (1966) on Ghanaian students to determine the ability to interpret pictorial depth the researchers came to the following conclusion:
Perceptual processing among different cultural groups are not clear cut. "The hypothesis that cultural stimulus is critical for the development of pictorial depth perception is supported by the present results. Surveys undertaken in the communities and homes of all the children studied revealed no evidence of activities such as reading, drawing, painting, looking at pictures, pattern-making or playing with construction toys, and it was exceptional for a child to have used a pencil prior to going to school. ... The opportunity for informal pictorial experience was, therefore, negligible" (Mundy-Castle, p.129, 1966, cf. Glick, 1984).

There doesn't appear to be enough evidence and very little research about the factors that determine perceptual processing (Glick, 1984). The best conclusion suggested that any specific cross-cultural comparisons, without taking into account etiological factors are premature (Glick, 1984).
CHAPTER III

METHODOLOGY

This chapter contains a description of the research setting, sample selection, procedures employed for collecting the data and the instruments utilized in the study. These sections are followed by a description of the statistics to be used in the data analysis.

Research Setting

The setting for this research were schools in the metropolitan area of Cape Town, South Africa. The South African education system is divided into four different provincial education systems (Cape, Transvaal, Orange-Free state and Natal). The Cape Province was chosen for the present study.

The Cape Province is at the southern corner of South Africa with a coastal zone bordering the Atlantic and the Indian Ocean. Cape Town is commonly referred to as the mother city, with the infamous table mountain landmark and is regarded as one of the most beautiful cities in the world. Table mountain is clearly visible over the city of Cape Town and Table Bay, north of the Cape Peninsula (Joyce & Barker, 1994).
Greater Cape Town area is covered with nearly 187,000 hectares (1 hectare is a metric unit equal to 2.471 acres) of area and has a population of approximately three million, which is 74.8% Blacks, 14.1% Whites, 8.5% Coloured, and 2.6% Indian (Asians) (South African Foundation, 1993).

Sample selection

Students selected for the present study were eight-and-nine-year-old school children currently enrolled in third (N = 200) and fourth (N = 200) grade classes. All four racial groups were represented in the present study. However, the investigator was unable to have access to white females schools. Both state and private schools were used for this study. This sample size was considered adequate for the statistical procedures to be applied to the data. School principal approval via consent form was obtained for each school before inclusion in the study (see Appendix A and B). In South Africa permission to conduct research within the schools and to gain access to the children is granted by the school principals and therefore, parental permission was not required.

Children who participated in the study were from working to upper-middle class backgrounds. In South Africa, private school education is accessible to only middle and upper-middle class children, and is comparable to standards of education in the United States. Hollingshead’s Index was used to verify socio-economic standing.
Types of Schools in South Africa

Private Schools

The following summary of the role of private schools in South Africa (J. Pampallis) was taken from a book by E. Unterhalter, H. Wolpe, and T. Botha (Eds, 1992), Education in a future South Africa: Policy Issues for transformation. Private schools serves mainly children of wealthy parents who can afford their fees to ensure that their children are educated in an atmosphere which they choose. These schools are extremely well equipped and well staffed. In many instances they supercede the state funded schools. They provide a good education for a minority. Private schools are generally English-speaking. There are many reasons why parents send their children to these schools. Some reasons are high academic standards, the use of innovative and progressive teaching methods, their philosophical or cultural values (perceived freedom from country’s ideological influences) and the opportunity for their children to associate and understand children from diverse racial and ethnic backgrounds. Thus these schools create a sense of eliteness within the consciousness of their students.

State Schools

Education in South Africa is mainly funded by the government. The education system, prior to the 1994 democratic election was characterized by severe racial disparities. For example, educational per capita expenditure: for a white child was U.S. $ 690, for an Indian child U.S. $430, for a coloured child U.S.
$295 and for a black child was U.S. $56.50 (Ormond, 1985). These meagre resources combined with under-qualified teaching staff characterized black schools. Another major problem lies in the disparity between teacher:pupil ratio. For example, White 1:20, Indian 1:23, Coloured 1:25 and Black 1:42 (SAIRR, 1989). Language also had an adverse affect on the education system. White, Indian and coloured children were taught in English and Afrikaans, whereas, the blacks were taught in the mother tongue language (i.e. Xhosa, Zulu, etc). English and Afrikaans were taught as either second or third languages. The aforementioned is a brief summary of the complexity of the education system in South Africa and how it catered to the different racial groups during the apartheid era.
Table 1

<table>
<thead>
<tr>
<th>Group</th>
<th>Grade</th>
<th>Age</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>*White</td>
<td>3</td>
<td>8</td>
<td>50</td>
<td></td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>9</td>
<td>50</td>
<td></td>
<td>50</td>
</tr>
<tr>
<td>Indian</td>
<td>3</td>
<td>8</td>
<td>25</td>
<td>25</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>9</td>
<td>25</td>
<td>25</td>
<td>50</td>
</tr>
<tr>
<td>Coloured</td>
<td>3</td>
<td>8</td>
<td>25</td>
<td>25</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>9</td>
<td>26</td>
<td>24</td>
<td>50</td>
</tr>
<tr>
<td>Blacks</td>
<td>3</td>
<td>8</td>
<td>27</td>
<td>23</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>9</td>
<td>26</td>
<td>24</td>
<td>50</td>
</tr>
<tr>
<td>Sample Size</td>
<td></td>
<td></td>
<td>254</td>
<td>146</td>
<td>400</td>
</tr>
</tbody>
</table>

*Note: White group has no females.

Procedures

Permission to conduct this research study was obtained from the participating school principals in Cape Town, South Africa. Before the beginning of each testing session, the investigator met with the school principal who formally introduced the investigator to the children. The following was an *oral solicitation to the students:
“My name is Andrew Martin and I am a student at The Ohio State University in the United States of America. I am here to ask for your help in a very important project. I am trying to learn how students do some different activities. Over the next few weeks, I will be working with students at different schools. With these activities you will be asked to solve some problems and do some puzzle activities. Most of the students find them interesting and fun to do. The total time required to do all these activities will be just over an hour. You will not be graded on any of these activities.”

The testing sessions involved from 30 to 50 pupils each. The testing was administered by the investigator with the presence of the teachers of those particular grades. The testing of the children occurred in regular classrooms and in some instances the testing was conducted in the school auditoriums. All the rooms were well lit, and seating arrangements were designed for the comfort of eight and nine years old, therefore, small tables and chairs were utilized in the testing rooms. The NNAT consists of 38 multiple choice items. The length of time required to complete the test was 30 minutes. The instructions for the NNAT and SAT tests were read aloud to students as they were asked to follow the written questions in the test booklet. The NNAT was first administered. Before the testing session commenced the students completed personal data on the NNAT answer sheet. While the students provided this information, the investigator answered definition questions as they arose.

Note: An (*) asterisk indicates that the examiner had to translate the oral solicitation in some instances into two other languages, Afrikaans and Xhosa.
Column G on the NNAT answer sheet has been changed in order to make the ethnicity section more applicable to South Africa. Column F, H and J were not used in this study. The following are samples of the **NNAT Answer Sheet and NNAT Sample Items:

![NNAT Answer Sheet]

Figure 1: Naglieri Nonverbal Ability Test Answer Sheet

**From Naglieri Nonverbal Ability Test: Tryout Edition. Copyright 1995 by The Psychological Corporation. Reproduced by permission. All rights reserved.**
Figure 2: Naglieri Nonverbal Ability Test Sample Items

**From Naglieri Nonverbal Ability Test: Tryout Edition. Copyright 1995 by The Psychological Corporation. Reproduced by permission. All rights reserved.
After all students completed the NNAT, the SAT were administered in a group format as well. Before the administration of the SAT the students were given a 10 - 15 minutes break. Since the students had provided the personal information on the NNAT, the SAT answer sheet was photocopied on the back of the NNAT answer sheet. The order of administration was Reading vocabulary first and then Concepts of number. The amount of time allocated to complete the Reading Vocabulary subtest was 30 minutes. Upon the completion of this task the students were instructed to continue with the Concepts of Number subtest and to complete this task within 25 minutes. Following on pages 42 and 43 are samples of the Stanford Achievement Test Answer Sheet and Sample Items of the two subtests.
Combination Answer Folder
Basic Battery/Science and Social Science

STANFORD
Achievement Test

with
OLSAT
Otis-Lennon School Ability Test

Reading Vocabulary

<table>
<thead>
<tr>
<th>SAMPLE A</th>
<th>SAMPLE B</th>
<th>SAMPLE C</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 3 2 1</td>
<td>13 8 5 2</td>
<td>21 3 2 1</td>
</tr>
<tr>
<td>6 2 1 3</td>
<td>14 3 2 1</td>
<td>22 1 3 2</td>
</tr>
<tr>
<td>7 1 3 2</td>
<td>15 2 3 1</td>
<td>23 1 2 3</td>
</tr>
<tr>
<td>8 1 2 3</td>
<td>16 3 2 1</td>
<td>24 2 1 3</td>
</tr>
<tr>
<td>9 1 2 3</td>
<td>17 3 2 1</td>
<td>25 3 1 2</td>
</tr>
<tr>
<td>10 3 1 2</td>
<td>18 2 3 1</td>
<td>26 1 2 3</td>
</tr>
<tr>
<td>11 3 1 2</td>
<td>19 2 3 1</td>
<td>27 3 1 2</td>
</tr>
<tr>
<td>12 2 1 3</td>
<td>20 3 1 2</td>
<td>28 1 2 3</td>
</tr>
</tbody>
</table>

Concepts of Number

<table>
<thead>
<tr>
<th>SAMPLE</th>
<th>SAMPLE</th>
<th>SAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 5</td>
<td>7 2 1 3</td>
<td>14 1 2 3</td>
</tr>
<tr>
<td>2 3 1 5</td>
<td>8 2 1 3</td>
<td>15 1 2 3</td>
</tr>
<tr>
<td>3 1 2 5</td>
<td>9 2 1 3</td>
<td>16 1 2 3</td>
</tr>
<tr>
<td>4 1 2 3</td>
<td>10 2 1 3</td>
<td>17 2 1 3</td>
</tr>
<tr>
<td>5 1 2 3</td>
<td>11 2 1 3</td>
<td>18 2 1 3</td>
</tr>
<tr>
<td>6 1 2 3</td>
<td>12 2 1 3</td>
<td>19 2 1 3</td>
</tr>
<tr>
<td></td>
<td>13 2 1 3</td>
<td>20 2 1 3</td>
</tr>
</tbody>
</table>

Figure 3: Stanford Achievement Test Answer Sheet

***From Stanford Achievement Test- 8th Edition. Copyright 1988 by The Psychological Corporation. Reproduced by permission. All rights reserved.
Sample Items of The Stanford Achievement Reading Vocabulary Subtest

A). Something that is huge is very-
   a) damp          c) pretty
   b) big           d) bright

B). Because the child was very cautious, he looked both ways before crossing the street. Cautious means-
   a) happy          c) playful
   b) silly          d) careful

C). He had a ring on his finger.
   In which sentence does the word ring mean the same as in the sentence above?
   a) He lost his new key ring.
   b) The teacher will ring the bell.
   c) The children held hands to form a ring
   d) She was wearing a gold ring.

Sample Item of The Stanford Achievement Concepts of Number Subtest

A. What number is one more than three?
   a) 4   b) 1   c) 2   d) 5

***Sample items of both subtests from Stanford Achievement Test- 8th Edition. Copyright 1988 by The Psychological Corporation. Reproduced by permission. All rights reserved.
No problems were experienced during the data collection. The data collection took place over a period of five weeks (August 1 through September 1, 1995).

**Instrumentation**

Two instruments were employed in this investigation: Naglieri Nonverbal Ability Test (NNAT) and Stanford Achievement Test (SAT; two subtests, Reading Vocabulary and Concepts of Number. Each instrument is described below:

Naglieri Nonverbal Ability Test: The Naglieri Nonverbal Ability Test (NNAT) is comprised of 38 multiple choice (5 options) items and is intended to be group administered. The test is intended to be a general nonverbal measure of intelligence like other tests that utilize the standard progressive matrix format (Naglieri, 1985). The NNAT is composed of items of the types found in the Ravens Progressive Matrices (Carlson and Wiedl, 1976). The test includes items in groups called Pattern Completion (requires the individual to choose the options which accurately completes the pattern), Reasoning by Analogy (requires the examinee to see how the change or changes in one figure relates to the analogous change or changes in another), Serial Reasoning (requires the individual to discover the order in which items appear throughout the matrix) and Spatial Visualization (requires the individual to imagine how a figure would look like when two or more designs are combined). The NNAT can be
administered to ages 5 to 17 to years old. The test is useful with linguistically and cultural diverse populations, since the administration requires minimal verbal directions. Several psychometric investigations have been conducted to support the theoretical properties and usefulness of the instrument in applied settings (Naglieri & Bardos, 1988; Haddad & Juliano, 1991).

Internal reliability for this instrument was established by calculating Cronbach's Alpha for each standardization age group using raw scores (Naglieri, 1985). The age range of the students in the standardization sample was 5-17. Internal reliability coefficients ranged from .63 (age 5.6 -5.11) to .89 (ages 8.0 -8.5 and 9.6 -9.11), with a median coefficient being .83 (Naglieri, 1985). The manual noted that all test-retest reliability coefficients for the MAT were obtained on a sample tested four weeks after the initial testing. Correlation coefficients computed by student grade in order to minimize ability related inflation were found to range between .51 and .91. The median test-retest coefficient across all grades was .75. The standard error of measurement ranged from 1.8 - 2.4 depending upon the age of the individual.

Developmental changes in mean scores were identified and item factor analysis were performed for each grade level of the standardization sample in order to determine the construct validity of the MAT. A clear increase in raw score means with age, which demonstrates age differentiation for the MAT was found. The results of item analysis revealed that most of the items for grades 2, 5 and 9 were above .40.
Concurrent validity was established by correlating the MAT with achievement measures such as Comprehensive Test of Basic Skills (CTBS; McGraw-Hill, 1969) and significant correlations were obtained between the MAT and CTBS composite scores in Reading ($r = .42, p < .01$), Spelling ($r = .29, p < .05$), Language ($r = .49, p < .01$), Math ($r = .39, p < .01$) and with the Total Battery ($r = .40, p < .01$). This finding is consistent with the significant correlations reported in the MAT manual (Naglieri, 1985a) between the MAT and the short form of the MAST's (MAST; Howell, Zucker & Morehead, 1985) Reading ($rs = .46$ and .53) and Mathematics scores ($rs = .44$ and .49) for normal 4th and 5th grade students. Other predictive research studies reported significant results. Smith (1988) administered the MAT and the Kaufman Assessment Battery for Children (K-ABC) to handicapped second graders and gifted students. Smith (1988) obtained a correlation between the MAT and K-ABC Mental Processing Composite of .53 for the handicapped subjects; among the same group, the correlation between the MAT and the K-ABC Simultaneous Processing was .49. For the gifted group the same correlations were found (i.e. .53 and .49 respectively). Haddad & Juliano (1991) correlated the MAT with the Iowa Test of Basic Skills (ITBS; Hieronymous & Hoover, 1990) with 82 normal 4th grade students of low socioeconomic status, and reported correlations between the MAT and the ITBS of .51. It can be concluded from the aforementioned studies that there appears to be a clear relationship between the MAT and academic achievement.
Stanford Achievement Test

The Stanford Achievement Test - 8th Edition (SAT; 1988) is a nonverbal measure of achievement and can be easily group administered from Kindergarten through 12th grade. The test's first edition was published in 1923. Each revision was carefully undertaken to provide the following: (1) That the test content and the instructional goals are more or less the same; (2) That the norms are an accurate reflection of the grade level; and (3) that the test provide educators and parents with meaningful information. The Stanford Achievement Test is therefore a well established achievement instrument and has been widely used throughout the United States. The degree of reliability of the subtests of the Stanford Achievement Test is represented by three different approaches: (1) The Internal consistency method, using the Kruder-Richardson (KR-20) - it yields consistency within results from one test administration to another; (2) Alternate forms method- this method measures the precision of the instrument, since it takes into account difference that arises from the different testing situation and from the different but similar test content, and (3) The consistency-over-time method. The first two is most useful in a reliability review.

The Stanford Achievement Test - 8th Edition (1988) yields grade equivalent scores and scaled scores based on the performance of both normal- and hearing-impaired children. The median internal consistency reliability coefficients
Cronbach's alpha were reported for the following subtests: Reading Vocabulary (.92) and Concepts of Number (.86). The Alternate reliability coefficients were also reported, Reading Vocabulary (X = 25; SD = 8.1) and Concepts of Number (X = 22.2; SD = 6.8). Following is a brief description of the two subtests that has been administered:

**Reading Vocabulary**

The reading vocabulary subtest takes the developmental stage of the child into consideration. Therefore the textual content of every subtest reflects the appropriate age level. The content of the reading vocabulary subtests uses three modes: Synonyms/definitional phrases, contextual clues, and multiple-meaning words. All the items in this subtest were presented in a multiple choice format. The main objective of this subtest therefore, is to measure two aspects. First, to assess the student's level of cognitive functioning. And, secondly, to assess the student's repertoire.

**Concepts of Number**

The Concepts of number subtest assess's the student's understanding mathematical skills. The concepts of number subtest tasks which ranges from adding and being able to read numerals in order to estimate and solving fractions. All the items in this subtest were presented in a multiple choice format.
Data Analysis

The data will be analyzed in accordance with the research questions presented in Chapter I. The methods of analysis employed to answer these questions will follow:

1. Performance of Black, White, Coloured and Indian students on the NNAT:
   a. Is there a significant difference between groups, by gender within each group?
   b. Is there a significant difference between groups, by race? Question 1a will be analyzed by using an independent t-Test (to determine the difference between uncorrelated means). Question 1b will be analyzed by using Analysis of Variance (ANOVA) to determine if a significance difference exist.

2. Performance of Black, White, Coloured and Indian students on the SAT:
   a. Is there a significant difference between groups, by gender within each group?
   b. Is there a significant difference between groups, by race? Question 2a will be analyzed by using an independent t-Test (to determine the difference between uncorrelated means). Question 2b will be analyzed by using Analysis of Variance (ANOVA) to determine if a significant difference exist.
3. Overall relationship between the NNAT and SAT:
   a. What are the relationships between the NNAT and SAT for the entire sample and by sex or race? This question will be analyzed by using a Pearson Product-Moment Correlation

Additional Analyses

4. Is there a significant difference between groups on NNAT by socioeconomic status (SES)?

5. Is there a significant difference between groups on the Stanford Achievement Concepts of number subtest, by socioeconomic status (SES)?

6. Is there a significant difference between groups on the Stanford Achievement Reading Vocabulary subtest, by socioeconomic status (SES)?

Questions 4, 5 and 6 will be analyzed by using Analysis of Variance to determine if significant differences exist.

7. Factorial Analysis of Variance of NNAT by Race and Socioeconomic Status (SES).

8. Factorial Analysis of Variance of Stanford Achievement Concepts of Number subtest by Race and Socioeconomic Status (SES).

9. Factorial Analysis of Variance of Stanford Achievement Reading Vocabulary subtest by Race and Socioeconomic Status (SES).
CHAPTER IV

RESULTS

The purpose of this study was to determine how the different racial
groups of South Africa perform on two measures of nonverbal assessment. The
specific aim of this study was to examine the predictive validity of the NNAT
using a group administered achievement test (SAT) as a criterion. This chapter
contains the findings of the study.

Sample Demographics

The sample consisted of 400 subjects in grades 3 (50%) and 4 (50%)
equally divided between ages eight and nine years. The sample consisted of
254 males (63.5%) and 146 females (36.5%), and the racial composition of this
sample included 100 Caucasians (25%), 100 Indians (25%), 100 Coloured
(25%), and 100 Blacks (25%). The socioeconomic status of the participants'
families varied in that 208 were professional (52%), 67 semi-skilled (16.75%),
and 125 unskilled (31.25%). Frequencies and percentages for the demographic
variables are presented in Table 2.
Table 2

Sample Demographics by Frequency and Percent*

<table>
<thead>
<tr>
<th>Variable</th>
<th>White</th>
<th>Indian</th>
<th>Coloured</th>
<th>Black</th>
<th>f</th>
<th>%</th>
<th>Cum%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>200</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>9</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>200</td>
<td>50</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>400</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>400</td>
<td>100</td>
<td>100.0</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>100</td>
<td>50</td>
<td>51</td>
<td>53</td>
<td>254</td>
<td>63.5</td>
<td>63.5</td>
</tr>
<tr>
<td>Female</td>
<td>*****</td>
<td>50</td>
<td>49</td>
<td>47</td>
<td>146</td>
<td>36.5</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>400</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>400</td>
<td>100</td>
<td>100.0</td>
</tr>
<tr>
<td>SES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unskilled</td>
<td>*****</td>
<td>*****</td>
<td>46</td>
<td>79</td>
<td>125</td>
<td>31.25</td>
<td>31.25</td>
</tr>
<tr>
<td>Clerical</td>
<td>*****</td>
<td>*****</td>
<td>49</td>
<td>18</td>
<td>67</td>
<td>16.75</td>
<td>48.0</td>
</tr>
<tr>
<td>Professional</td>
<td>100</td>
<td>100</td>
<td>5</td>
<td>3</td>
<td>208</td>
<td>52.0</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td>400</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>400</td>
<td>100</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Note: No females in white group.

*All variables were examiner's identified except SES which was based on parent self-report.

Table 3 presents the t-test results of the South African sample by gender on the NNAT and selected Stanford Achievement subtests. This analysis is summarized in Table 3 on page 53. An examination of Table 3 revealed that no statistical significant differences were found between genders within each group.
### Table 3

<table>
<thead>
<tr>
<th>Group</th>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>t</th>
<th>df</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>NNAT</td>
<td>100</td>
<td>28.26</td>
<td>5.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SNUM</td>
<td></td>
<td>26.51</td>
<td>4.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SREAD</td>
<td></td>
<td>27.34</td>
<td>4.58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indian</td>
<td>NNAT</td>
<td>50</td>
<td>25.10</td>
<td>6.85</td>
<td>50</td>
<td>24.54</td>
<td>6.89</td>
<td>.4075</td>
<td>98</td>
<td>.684</td>
</tr>
<tr>
<td></td>
<td>SNUM</td>
<td></td>
<td>24.66</td>
<td>5.87</td>
<td></td>
<td>24.22</td>
<td>6.04</td>
<td>.3691</td>
<td>98</td>
<td>.712</td>
</tr>
<tr>
<td></td>
<td>SREAD</td>
<td></td>
<td>23.96</td>
<td>3.31</td>
<td></td>
<td>23.90</td>
<td>3.12</td>
<td>.0932</td>
<td>98</td>
<td>.926</td>
</tr>
<tr>
<td>Coloured</td>
<td>NNAT</td>
<td>51</td>
<td>20.96</td>
<td>6.55</td>
<td></td>
<td>21.16</td>
<td>6.48</td>
<td>.1552</td>
<td>98</td>
<td>.877</td>
</tr>
<tr>
<td></td>
<td>SNUM</td>
<td></td>
<td>20.74</td>
<td>6.15</td>
<td></td>
<td>20.55</td>
<td>5.92</td>
<td>.1605</td>
<td>98</td>
<td>.872</td>
</tr>
<tr>
<td></td>
<td>SREAD</td>
<td></td>
<td>22.86</td>
<td>6.06</td>
<td></td>
<td>22.38</td>
<td>6.37</td>
<td>.3817</td>
<td>98</td>
<td>.703</td>
</tr>
<tr>
<td>Blacks</td>
<td>NNAT</td>
<td>53</td>
<td>17.79</td>
<td>5.65</td>
<td>47</td>
<td>16.61</td>
<td>4.58</td>
<td>1.131</td>
<td>98</td>
<td>.260</td>
</tr>
<tr>
<td></td>
<td>SNUM</td>
<td></td>
<td>17.50</td>
<td>7.19</td>
<td></td>
<td>16.63</td>
<td>5.28</td>
<td>.6829</td>
<td>98</td>
<td>.496</td>
</tr>
<tr>
<td></td>
<td>SREAD</td>
<td></td>
<td>18.96</td>
<td>7.19</td>
<td></td>
<td>18.00</td>
<td>4.69</td>
<td>.7806</td>
<td>98</td>
<td>.436</td>
</tr>
</tbody>
</table>

*** SNUM= Concepts of Number

*** SREAD= Reading Vocabulary
Table 4 presents the mean NNAT raw scores earned by the South African White, Indian, Coloured and Black students by group and grade level. An examination of Table 4 revealed a clear increase in the mean raw scores of NNAT by grade, which demonstrates age differentiation for the NNAT (see Figure 4 on page 55).

Table 4

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Third Grade</th>
<th>N</th>
<th>Fourth Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>White</td>
<td>50</td>
<td>26.14</td>
<td>6.21</td>
<td>30.38</td>
</tr>
<tr>
<td>Indian</td>
<td>50</td>
<td>21.50</td>
<td>6.96</td>
<td>28.14</td>
</tr>
<tr>
<td>Coloured</td>
<td>50</td>
<td>17.24</td>
<td>6.72</td>
<td>24.88</td>
</tr>
<tr>
<td>Blacks</td>
<td>50</td>
<td>14.66</td>
<td>4.40</td>
<td>19.82</td>
</tr>
</tbody>
</table>

Table 5 on page 55 presents the means and standard deviations for the selected Stanford Achievement subtests by group and grade level. An examination of these mean raw scores, indicates that the mean raw scores of the subtests for the South African White, Indian, Coloured and Blacks students increased consistently with grade level (see Figures 5 and 6 on pages 57 and 58).
Figure 4: NNAT Raw Score Means by Grade and Race
Table 5

Means and Standard Deviations of the selected Stanford Achievement subtests by group and grade

<table>
<thead>
<tr>
<th>Group</th>
<th>Variable</th>
<th>N</th>
<th>Third Grade</th>
<th>N</th>
<th>Fourth Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>White</td>
<td>SNUM</td>
<td>50</td>
<td>23.90</td>
<td>4.38</td>
<td>29.12</td>
</tr>
<tr>
<td></td>
<td>SREAD</td>
<td>50</td>
<td>25.18</td>
<td>4.46</td>
<td>29.50</td>
</tr>
<tr>
<td>Indian</td>
<td>SNUM</td>
<td>50</td>
<td>21.30</td>
<td>5.98</td>
<td>27.58</td>
</tr>
<tr>
<td></td>
<td>SREAD</td>
<td>50</td>
<td>22.42</td>
<td>2.97</td>
<td>25.44</td>
</tr>
<tr>
<td>Coloured</td>
<td>SNUM</td>
<td>50</td>
<td>16.96</td>
<td>5.29</td>
<td>24.34</td>
</tr>
<tr>
<td></td>
<td>SREAD</td>
<td>50</td>
<td>19.02</td>
<td>5.44</td>
<td>26.24</td>
</tr>
<tr>
<td>Blacks</td>
<td>SNUM</td>
<td>50</td>
<td>14.68</td>
<td>5.68</td>
<td>19.52</td>
</tr>
<tr>
<td></td>
<td>SREAD</td>
<td>50</td>
<td>16.20</td>
<td>4.85</td>
<td>20.82</td>
</tr>
</tbody>
</table>
Figure 5: Stanford Achievement Reading Vocabulary subtest Raw Score Means by Grade and Race
Figure 6: Stanford Achievement Concepts of number subtest Raw Score Means by Grade and Race
A two-factor fixed-model Anova was performed to determine the significant treatment and interaction effects on NNAT scores. This analysis is summarized in Table 6. The results suggest two significant findings. The $F$ for race was statistically significant, $F (3, 392) = 80.91, p < .01$. The $F$ for the interaction of race and grade was not statistically significant, $F (3, 392) = 2.05, p > .01$.

Table 6

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td>3</td>
<td>2260.8366</td>
<td>80.91</td>
<td>.001</td>
</tr>
<tr>
<td>Grade</td>
<td>1</td>
<td>3504.6400</td>
<td>125.42</td>
<td>.001</td>
</tr>
<tr>
<td>Race X Grade</td>
<td>3</td>
<td>57.3066</td>
<td>2.05</td>
<td>.113</td>
</tr>
<tr>
<td>Error</td>
<td>392</td>
<td>27.9421</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>399</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A two-factor fixed-model Anova was performed to determine significant treatment and interaction effects on Stanford Achievement Concepts of Number subtest scores. This analysis is summarized in Table 7 on page 60. The results suggest two significant findings. The $F$ for race was statistically significant, $F (3, 392) = 72.65, p < .01$. The $F$ for the interaction of race and grade was not statistically significant, $F (3, 392) = 1.37, p > .01$. 
A two-factor fixed -model Anova was performed to determine significant treatment and interaction effects on Stanford Achievement Reading Vocabulary subtest scores. This analysis is summarized in Table 8 on page 61. The results suggest three significant findings. The $F$ for race was statistically significant, $F (3, 392) = 64.49$, $p < .01$. The $F$ for grade was also statistically significant, $F (1, 392) = 111.33$, $p < .01$. However, the most important finding is the $F$ for the interaction of race and grade, $F (3, 392) = 3.75$, $p < .01$. 

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td>3</td>
<td>1733.4566</td>
<td>72.65</td>
<td>.001</td>
</tr>
<tr>
<td>Grade</td>
<td>1</td>
<td>3516.4900</td>
<td>147.38</td>
<td>.001</td>
</tr>
<tr>
<td>Race X Grade</td>
<td>3</td>
<td>32.6433</td>
<td>1.37</td>
<td>.252</td>
</tr>
<tr>
<td>Error</td>
<td>392</td>
<td>23.8559</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>399</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 8

Analysis of Variance (ANOVA) of Stanford Achievement Reading Vocabulary subtest by Race and Grade

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td>3</td>
<td>1331.8491</td>
<td>64.49</td>
<td>.001</td>
</tr>
<tr>
<td>Grade</td>
<td>1</td>
<td>2299.2025</td>
<td>111.33</td>
<td>.001</td>
</tr>
<tr>
<td>Race X Grade</td>
<td>3</td>
<td>77.3958</td>
<td>3.75</td>
<td>.001</td>
</tr>
<tr>
<td>Error</td>
<td>392</td>
<td>20.6527</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>399</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 9 on pages 62 and 63 presents the Pearson-Product Moment Correlations for the different groups and the total sample. The results suggest that all correlations were statistically significant.

Table 9

Correlations among NNAT and selected Stanford Achievement subtests by group and Total Sample (N = 100 for each group).

<table>
<thead>
<tr>
<th></th>
<th>NNAT</th>
<th>SNUM</th>
<th>SREAD</th>
<th></th>
<th>NNAT</th>
<th>SNUM</th>
<th>SREAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td></td>
<td></td>
<td></td>
<td>Indian</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>1.00</td>
<td>.63*</td>
<td>.65*</td>
<td>White</td>
<td>1.00</td>
<td>.82*</td>
<td>.60*</td>
</tr>
<tr>
<td>NNAT</td>
<td>1.00</td>
<td>.73*</td>
<td></td>
<td>SNUM</td>
<td>1.00</td>
<td>.65*</td>
<td>SREAD</td>
</tr>
<tr>
<td>SNUM</td>
<td>1.00</td>
<td>.60*</td>
<td></td>
<td>SREAD</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SREAD</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coloured</th>
<th>NNAT</th>
<th>SNUM</th>
<th>SREAD</th>
<th></th>
<th>NNAT</th>
<th>SNUM</th>
<th>SREAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coloured</td>
<td>1.00</td>
<td>.83*</td>
<td>.85*</td>
<td></td>
<td>Blacks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NNAT</td>
<td>1.00</td>
<td>.88*</td>
<td></td>
<td></td>
<td>1.00</td>
<td>.70*</td>
<td>.79*</td>
</tr>
<tr>
<td>SNUM</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td>.82*</td>
<td></td>
</tr>
<tr>
<td>SREAD</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: An (*) asterisk indicates correlations were significant p < .001.
A one-way analysis of variance (ANOVA) was performed to determine the significance of socioeconomic status on the performance of NNAT. This analysis is summarized in Table 10. The results suggest a significant finding. The $F$ for SES was statistically significant, $F(2, 397) = 170.55, p < .01$.

Table 10

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>SES</td>
<td>2</td>
<td>49.476305</td>
<td>170.55</td>
<td>.001</td>
</tr>
<tr>
<td>Error</td>
<td>397</td>
<td>29.0104</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>399</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: An (*) asterisk indicates correlations were significant $p < .001$. 
A one-way analysis of variance (ANOVA) was performed to determine
the significance of socioeconomic status on the performance of Stanford
Achievement Concepts of Number subtest. This analysis is summarized in
Table 11. The results suggest a significant finding. The F for SES was
statistically significant, \( F (2, 397) = 154.18, p < .01 \).

Table 11

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>SES</td>
<td>2</td>
<td>3971.1756</td>
<td>154.18</td>
<td>.001</td>
</tr>
<tr>
<td>Error</td>
<td>397</td>
<td>25.7566</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>399</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A one-way analysis of variance (ANOVA) was performed to determine
the significance of socioeconomic status on the performance of Stanford
Achievement Reading Vocabulary subtest. This analysis is summarized in
Table 12 on page 65. The results suggest a significant finding. The F for SES
was statistically significant, \( F (2, 397) = 176.31, p < .01 \).
A factorial analysis of variance was performed to determine the main effect of race and socioeconomic status (SES) on the performance of NNAT and the interaction between them. This analysis is summarized in Table 13 on page 66. The results suggest the following findings. The $F$ for race was statistically significant, $F (3, 392) = 8.49, p < .01$. The $F$ for SES was also statistically significant, $F (2, 392) = 66.77, p < .01$. However, the interaction between race and socioeconomic status was not statistically significant, $F (2, 392) = 0.04, p < .01$. 

Table 12

**Analysis of Variance (ANOVA) of Stanford Achievement Reading Vocabulary subtest by Socioeconomic Status (SES)**

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>$F$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SES</td>
<td>2</td>
<td>3439.3026</td>
<td>176.31</td>
<td>.001</td>
</tr>
<tr>
<td>Error</td>
<td>397</td>
<td>19.5067</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>399</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 13

Factorial Analysis of Variance of NNAT by Race and Socioeconomic Status (SES)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td>3</td>
<td>233.8376</td>
<td>8.49</td>
<td>.0001</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1839.6626</td>
<td>66.77</td>
<td>.0001</td>
</tr>
<tr>
<td>Race X SES</td>
<td>2</td>
<td>0.9919</td>
<td>0.04</td>
<td>.9646</td>
</tr>
<tr>
<td>Error</td>
<td>392</td>
<td>27.5526</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>399</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-Square</td>
<td>.4955</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A factorial analysis of variance was performed to determine the main effect of race and socioeconomic status (SES) on the performance of the Stanford Achievement Concepts of Number subtest and the interaction between them. This analysis is summarized in Table 14 on page 67. The results suggest the following findings. The F for race was statistically significant, $F(3, 392) = 5.41$, $p < .01$. The F for SES was also statistically significant, $F(2, 392) = 65.56$, $p < .01$. However, the interaction between race and socioeconomic status was not statistically significant, $F(2, 392) = 2.29$, $p < .01$. 
Table 14

Factorial Analysis of Variance of Stanford Achievement Concepts of Number subtest by Race and Socioeconomic Status (SES)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td>3</td>
<td>133.5698</td>
<td>5.41</td>
<td>.0012</td>
</tr>
<tr>
<td>SES</td>
<td>2</td>
<td>1619.5341</td>
<td>65.56</td>
<td>.0001</td>
</tr>
<tr>
<td>Race X SES</td>
<td>2</td>
<td>56.4512</td>
<td>2.29</td>
<td>.1031</td>
</tr>
<tr>
<td>Error</td>
<td>392</td>
<td>24.7043</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>399</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>R-Square</strong></td>
<td></td>
<td>.4669</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A factorial analysis of variance was performed to determine the main effect of race and socioeconomic status (SES) on the performance of the Stanford Achievement Reading Vocabulary subtest and the interaction between them. This analysis is summarized in Table 15 on page 68. The results suggest the following findings. The F for race was statistically significant, $F(3, 392) = 17.16$, $p < .01$. The F for SES was also statistically significant, $F(2, 392) = 113.87$, $p < .01$. However, the interaction between race and socioeconomic status was not statistically significant, $F(2, 392) = 3.00$, $p < .01$. 

67
Table 15

Factorial Analysis of Variance of Stanford Achievement Reading Vocabulary subtest by Race and Socioeconomic Status (SES)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td>3</td>
<td>293.8467</td>
<td>17.16</td>
<td>.0001</td>
</tr>
<tr>
<td>SES</td>
<td>2</td>
<td>1949.9812</td>
<td>113.87</td>
<td>.0001</td>
</tr>
<tr>
<td>Race X SES</td>
<td>2</td>
<td>51.4075</td>
<td>3.00</td>
<td>.0508</td>
</tr>
<tr>
<td>Error</td>
<td>392</td>
<td>17.1248</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>399</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-Square</td>
<td>.5409</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER V

DISCUSSION

The primary purpose of the present study was to determine the performance of a diverse South African sample on the NNAT. A secondary aim was to predict the validity of the NNAT by using an achievement test (SAT) as a criterion.

In order to answer these questions, a sample of 400 students who attended state and private schools were selected for the study. The 400 subjects in this study represented variation in gender, ethnicity and socioeconomic (SES) level. The sample was equally distributed in terms of ethnicity, Caucasian (25%), Indian (25%), Coloured (25%) and Blacks (25%). However, 52% of the sample were represented by the children from professional families and 31.25% of the children were from families where the parents were unskilled laborers and 16.75% of the children were from clerical families.
Performance of Black, White, Coloured and Indian South African students on NNAT: Is there a significant difference between groups, by gender with each group?

In the present study, reported results of the NNAT within each group showed no statistically significant differences. Thus, gender did not play a significant role in the overall performances of the different racial groups.

Is there a significant difference between groups, by race?

In the present study, intergroup comparisons between the mean raw scores of the different racial groups reported results that displayed a statistical significant difference at third grade level. For example, intergroup comparison between the white South African sample of third grade students with black South African third grade students, reported a 11.4 point raw score difference. The present study also reported intergroup comparisons between white South African sample of third grade students with coloured South African third grade students, a 8.90 point raw score difference. The reported differences may be attributed to a number of factors. For example, researchers have reported that socioeconomic status (SES; Slaney & Brown, 1983) and other moderator variables such as age and gender (Johnson & Brems, 1990) account for more total variance than does any other factor. Other possible contributing factor, is the level of schooling, previous experience and educational background. It is not evident from the results whether each group was looking for
non-matching items, rather than matching ones. The only way we can ascertain
the answer is to employ Item Response Theory (IRT). However, the latter is
beyond the scope of the present study.

The present study also reported intergroup comparisons between the
white South African sample of the third grade students and Indian South
African third grade students, with a 5 point raw score difference. This
difference is small and not statistically significant. It is possible to infer that
this from this small variance is due to these subjects having more or less the
same socioeconomic status, education and family income etc.

In the present study, results of the NNAT reported that the fourth grade
students performed better than the third grades. Developmental changes in the
mean raw scores for each grade level were found. Developmental changes is a
basic aspect of construct validity (Anastasi, 1988). Thus, age differentiation was
established.

An examination of the intergroup comparisons of the fourth grades reported
statistically significant differences between the different groups. For example,
ingroup comparison between the white and black sample of fourth grade
students and between white and coloured, reported a 10.5 and 5.5 point
difference on the mean raw scores, respectively. In contrast, the intergroup
comparison between the white and indian sample, reported a 2 point raw score
difference on the mean raw scores.
No indication of different approaches to tasks requiring matching have been reported in the literature on abstract problem-solving ability of South African children. Further information on this dimension may be gathered by interviewing the subjects and determining the reasons for their chosen responses. Such a qualitative approach would determine the strategic applications of subjects to these nonverbal tasks. However, such an attempt is beyond the present study. Therefore, extensive research with a much larger sample would be necessary before conclusive statements may be made regarding matching skills of the South African children. No literature was found which examined item characteristic curves of South African children on nonverbal measures of intelligence and the use of such measures in cross-cultural research which had been adapted from Western culture. Furthermore, no cross-cultural research has yet been done with the NNAT, since it is newly developed. Thus no supportive statements could be made regarding the use of these type of intelligence tests. However, a noteworthy statement is that intelligence is known to be influenced by socioeconomic status (Jachuck, 1982; Narchal & Juneja, 1986); education, family size and income (Jixiu, 1985). Cultural differences, however, do not account for all variance. Thus potential differences in test scores may be due more to factors such as SES than ethnicity and biased instruments (Bond, 1990; Groth-Marnat, 1990). Thus intelligence may not be confined to race, but rather the individual’s social position.
Performance of Black, White, Coloured and Indian South African students on selected Stanford Achievement subtest:

Is there a significant difference between groups, by gender within each group?

In the present study, reported results of the Stanford Reading Vocabulary and Concepts of Number subtests within each racial group reported no statistical significant differences. An examination of the results on the performance of Reading Vocabulary between genders within each group reported results that showed no significant differences between genders. Similar results were reported for the Stanford Concepts of Number subtest. Thus, gender was not significant in the overall performance of these different racial groups.

Is there a significant difference between groups, by race?

In the present study, intergroup comparisons between the mean raw scores of the different racial groups reported results that displayed statistical significant findings. For example, intergroup comparisons between white South African third grade with black South African third grade students and white South African third grade students with coloured South African third grade students reported a 8.9 and 6.1 point raw score difference, respectively on the Reading Vocabulary subtest. Somewhat similar results (9.2 and 6.9 point raw score difference) were reported for the performance on the Concepts of Number for the same two intergroups comparisons. These reported results
may have been influenced by a number of important factors. The items on both these subtests were presented in a multiple choice format that included three incorrect and one correct answer. It was required from the child to create a strategy to solve the problem. The results showed that these racial groups arrived at different application approaches. These approaches, however, may be due to differences in educational environment, cultural strategic application of learning tools, etc. Thus, these differences will be reflected in academic achievement scores.

In the present study, results of the selected Stanford Achievement subtests reported that the fourth grade students performed better than the third grades. The results are consistent with the manual in that there should be developmental changes in the mean raw scores for each grade level. The reading vocabulary and the understanding of mathematical skills are more advanced for the fourth grades than the third grades. Thus, age differentiation was established. Despite the overall performance of this grade level on these two subtests, the results reported statistical differences between intergroup comparisons. For example, intergroup comparison between the white sample and the black sample and white and coloured sample reported a 8.6 and 3.2 point raw score difference, respectively on the Reading Vocabulary subtest. Similar results (9.6 and 4.7 point raw score difference, respectively) were reported for the same two intergroup comparisons on the Concepts of Number subtest. However, the intergroup comparison between the white and indian
fourth grades on both Reading Vocabulary and Concepts of Number subtests were (4.0 and 1.5 points raw score, respectively). The significant variance between the white and the black students may have been significantly contributed by the level of education and economic mobility. The latter may be an explanation for the small difference between the Indian and White sample. Overall, it is possible to infer that both the third and fourth grades possess problem-solving skills, as evidence of their abilities.

What are the relationships between the NNAT and SAT for the total sample and by race?

A Pearson Product Moment Correlation was employed in order to determine the relationship between the NNAT and SAT. Significant correlations were found between the NNAT and SAT for all the different racial groups. The strongest relationship was obtained for the black sample. A relationship of similar strength was obtained for the coloured sample. The observed correlation between the NNAT and SAT for the total sample was also significant. These correlations may be interpreted in many ways. For example, a significant relationship may be reflective of the degree that abstract and figural reasoning are linked to achievement. It appears, that an association may be influenced by a number of factors which is beyond the scope of the present study. Therefore, no conclusive statements may be made regarding the significance of NNAT and its relationship with SAT, since no
cross-cultural research has been done to examine this relationship.

Limitations of the Study and suggestions for Future research

One limitation of the present study was the unequal distribution of diverse economic status among the subjects. One-half of the subjects came from families who were professionals, thus tend to represent a somewhat skewed population in socioeconomic terms. Intelligence is known to be influenced by socioeconomic status (Jachuck, 1982; Narchal & Juneja, 1986); education, family size and income (Jixiu, 1985). This partially explains why the mean scores are lower for some groups in this sample. In addition to sampling, the fact that the subjects were not randomly selected is regarded as a limitation. In addition, the sampling population contained a high number (25%) who were from private schools and another 25% were from professional families. This may be have resulted in producing high mean raw scores for some groups. In addition, to sampling the fact that the Caucasian group had no females, can be regarded as a limitation of this study. Another limitation related to the population sample was the age range. The subjects were all eight and nine years of age. Future research should attempt to obtain a more heterogeneous population sample in terms of socioeconomic status and age range.

Another limitation of this study was that the findings can only be generalized to the participating schools in this study. The mere fact that all the subjects were from Cape Town, South Africa is a limitation of this study. Future
research should attempt to obtain a national sample that is randomly selected.

The testing format may have been problematic in that some of the subjects (especially blacks) had never been tested in multiple-choice format before. Thus no prior exposure to this type of format can be regarded as a limitation of this study. Future research should attempt to expose subjects to this type of format before engaging in any type of testing.

A time limit was placed on the completion of an instrument. Since the subjects were subjected to a specific time constraint to complete a test, there performance might have been adversely affected by the time limit. Timing the subject was therefore, regarded as a limitation of the study.

Related to instrumentation: Is the NNAT so "culture-fair" that it has similar factorial structures for subjects from a wide range of dissimilar backgrounds? Nonverbal tests are considered to have figural, abstract or low verbal content, thus it is inherently biased and therefore, constitute a limitation of the study. Future research should attempt to expose subjects to these type of material before engaging in any testing.

Related to the above limitation regarding instrumentation: Although the publisher states that the Stanford Achievement Test is a nonverbal test, in practice subjects are required to have reading skills in order to understand the content of the test. Thus the Stanford Achievement Test is regarded as a limitation, because it has verbal content. Related to the Stanford Achievement
Test, Reading Vocabulary subtest is based on the American school curriculum, thus it is a limitation of this study. Future research should attempt to based a Reading Vocabulary subtest on an appropriate culture’s school curriculum, to make a test more relevant. Related to the Reading Vocabulary subtest, the content of this subtest is verbal, therefore, cannot be regarded as a nonverbal achievement subtest. Thus it is a limitation of this study.

Conclusion

Related to the outcome of the present study, we can observe that one group performs better than another group. For each group in the study, student scores on the NNAT and SAT covaried, indicating positive and significant correlations. It is important to note, however, that correlations are not necessarily predictive and for these students the school and social conditions that produced higher achievement in reading and math might also be those that contributed to the higher performance scores on the NNAT. That is, scores on the NNAT as with the SAT may be mostly of a function of the enriched learning and nurturing conditions that typically accompany higher socioeconomic levels. If this is the case, than it may be possible to use socioeconomic status as a predictive measure of academic achievement.

Presumably instruments such as the NNAT are used to predict how well students will perform academically over an extended period of time. A longitudinal study where students are assessed on the NNAT prior to formal
schooling might be a more viable or valid assessment for this purpose. Nevertheless, it will never be possible to rule out the influence of culture, and the fact that group differences are obtained points to the ubiquitous role of culture. Thus, these differences in cognition may be due to differences in educational environment, cultural differences in the applications of learning tools, and access to learning opportunities as dictated by socioeconomic status.

The assumption may not be made that since NNAT is devoid of expressive language and utilizes figural and abstract thinking, that it is "culture-fair". The enthusiastic proponents of "culture-fair or culture-reduced" tests are incognizant to assume that perceptual, learning and reasoning processes are a-cultural (Jensen, 1980; Naglieri & Bardos, 1988; Raven, 1947). To assume that "culture-fair" tests are a-cultural is to injure the power of differences between people of different cultures. The latter is a characteristic of a philistine. In reality, perceptual, learning and reasoning processes are best understood within a culture's distinct frame of reference. Furthermore, these processes are influenced by the interactions which, in turn, determine the skills that are being learned and these learning skills are controlled by the cultural group's values and social systems. Culture equals thinking (Vygotsky, 1987) and is manifested in social praxis. Thus, culture plays a vital role in an individual's cognitive development. Therefore, those who wish to use and interpret measures of nonverbal assessment within a cross-cultural setting which has been standardized and developed within a western society should approach
these measures with caution.

The findings from the present study raise a critical issue relative to how instruments such as the NNAT might be used. For example, scores showing that individuals from depressed and low socioeconomic groups perform less well might be used to further limit learning and achievement opportunities. A related question pertains to the real need for such predictive measures. Except for the one or two percent at the extremes, there is little need for assessment measures that purport to indicate how well a student will perform in school, this is particularly true for those from economically and socially oppressed groups. Although much has been written about testing and the use of it, for these measures there appears to be no *a priori* theoretical underpinning regarding the development of intellectual capacity or the distribution of intelligence.

The Ravens Progressive Matrices, Test of Nonverbal Intelligence, Hiskey-Nebraska, Leiter International Performance Scale, Matrix Analogies Test, and other reported nonverbal tests have been promoted to measure nonverbal cognitive ability (Llabre, 1984; Matey, 1984; McMorris, 1989). A current comparable proclamation of such an endeavor is the NNAT. Except for updated norms, this test does not appear to expand on what has been done before in the area of nonverbal cognitive assessment. "Merely reorganizing or restating what is already known and what has already been written is not research, valuable as it might be as a learning experience. It adds nothing to
what is known” (Best, p. 9, 1970).

If research endeavors follow over traveled paths, then what place do such pursuits have in the profession of education and psychology? Exploration, creativity, and uniqueness are imperative qualities in any research endeavor regardless of how small or large the pursuit. Given the current trends in education and psychology, the fields can afford nothing less.

The results of the present study indicate that the NNAT correlates well with standardized instruments of achievement such as SAT. Whether or not the NNAT is an innate measure of intelligence cannot be answered by these data. Even less clear is the matter of whether this instrument is "culture-free". Such a proclamation cannot be taken lightly. In scientific investigations, “the scientist must always subject his notions to the court of empirical inquiry and test” (Kerlinger, p. 11, 1986). As previously explained, cross-cultural assessment is a complex process that has yet to be adequately and vigorously tackled. In order to make such a proclamation of “culture-free”, at the very least, one must present data from numerous cultures and then limit the findings to these cultures only. Group differences in the present study suggest otherwise, and past experiences point to the need for extreme caution since applications of instruments such as these often lead to further subjugate rather than affirm fragile populations. Thus, the NNAT fails to move beyond what has already been done before in the area of nonverbal cognitive assessment.
As Kerlinger (1986) aptly points out, "The scientist's job is to discover new facts and to add them to the already existing body of information" (p. 7).

Finally, in the spirit of making scientific and scholarly contributions to the helping fields of education and psychology, the focus must be on a genuine interest in helping people rather than on the mere action of promoting recycled old ideas. The research question that the NNAT must now answer is:

"So what?"


89


APPENDIX A
Oral Solicitation to the School Principal

My name is Andrew Martin and I am a student at The Ohio State University in the United States of America. I am here to ask for your help in a very important project. This research is being conducted under the supervision of Antoinette Miranda, Ph. D. I am trying to learn how students do some different activities. These activities involve some problems and some puzzle activities. Most of the students find them interesting and fun to do. Here is a copy of the instruments that I will be using in this study. The total time required to do all of this will be 80 minutes. The students will not be graded on any of these activities. The students do not have to participate in the project if they do not want to. It will not count against them if they do not want to participate. Also, if they decide to participate but change their mind, that is O.K., too. However, I think they will find it interesting and fun.

Note: (*) An asterisk indicates that this letter was adapted from Mary Dixon’s dissertation (1994).
APPENDIX B
*Letter to the School Principal

I am a Ph. D. candidate in the School Psychology Program at The Ohio State University in the United States of America. For my dissertation research, I am examining the way children solve nonverbal problems. This research is being conducted under the supervision of Antoinette Miranda, Ph. D. It is hoped that the results of this study may provide insight into the process involved in children’s thinking. This would provide important instructional information for educators. I would greatly appreciate the assistance of you and your school in this effort.

All students in third and fourth grade at your school will be asked to participate in this study. Their names will not be revealed in the study. Each student’s name who participate in the study will correspond with an identity number. The final disposition of the list of names will be discarded upon completion of the study. The student’s participation in the study would involve one testing session of approximately 80 minutes. The session need to occur at school, during the regular school day. During the session, they will be given problem solving tasks. These tasks are game-like activities which most children find interesting and fun and will be group administered to students by myself. I will make alternative arrangements with the respected teachers for non-participating students to be engaged in a planned activity. If you give your permission the students participation is still voluntary and he or she can withdraw at any time without any consequences.

Thank you for your assistance.

Sincerely

Andrew Martin
Doctoral Candidate

Antoinette Miranda, Ph. D.
Advisor, School Psychology

Note: An (*) asterisk indicates that this letter was adapted from Mary Dixon’s dissertation (1994).