28033

NAME OF AUTHOR/*NOM DE L'AUTEUR* John B. Acheson

TITLE OF THESIS/*TITRE DE LA THÈSE* The Standard Progressive Matrices In A Multi-Ethnic Milieu

UNIVERSITY/*UNIVERSITÉ* University of Alberta

DEGREE FOR WHICH THESIS WAS PRESENTED/*GRADE POUR LEQUEL CETTE THÈSE FUT PRÉSENTÉE* Ph.D

YEAR THIS DEGREE CONFERRED/*ANNÉE D'OBTENTION DE CE GRADE* 1975

NAME OF SUPERVISOR/*NOM DU DIRECTEUR DE THÈSE* E. E. Fox

DATED/*DATÉ* June 26/75 SIGNED/*SIGNÉ* J. Acheson

PERMANENT ADDRESS/*RÉSIDENCE FIXE* 4803- Malmo Rd
Edmonton Alberta

THE UNIVERSITY OF ALBERTA

THE STANDARD PROGRESSIVE MATRICES IN A MULTI-ETHNIC MILIEU

by

© JOHN B. ACHESON

A THESIS

SUBMITTED TO

THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

IN

EDUCATIONAL PSYCHOLOGY

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

EDMONTON, ALBERTA

FALL, 1975

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled "The Standard Progressive Matrices In A Multi-Ethnic Milieu" submitted by John B. Acheson in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

T. E. Fox
Supervisor

V. R. Nyberg

G. W. Fitzsimmons

Bernard Schwartz

J. E. Bicknell
External Examiner

Date Jan. 20, 1975

ABSTRACT

The objective of the present study was to develop a local set of norms for the Standard Progressive Matrices and to compare these norms to those developed by Raven and reported in the test manual. Concomitantly, tests of reliability, construct validity, and concurrent validity were carried out. Finally the Progressive Matrices was examined in both a varied socio-economic and cross-cultural context.

Seven hundred and thirty-six twelve, thirteen and fourteen year old students were randomly selected from the junior high school grades in the Edmonton Separate School System. An additional twenty-eight twelve and fourteen year old students were selected from grades six and ten respectively. A sample of 106 Canadian Indian students was drawn from the same population for the cross-cultural studies. Grade and age norms expressed in percentiles were developed plus a table of norms expressed as deviation intelligence scores.

As hypothesized, the norms developed in the present study differed significantly from those developed by Raven. A Kolmogorov-Smirnov Test showed the difference to be significant beyond the .001 level for the grade seven and eight students and beyond the .01 level for the grade nine students. In each case the students in the current

study scored significantly higher than the students in Raven's study.

Using the KR-20 formula, internal consistency measures of 0.86, 0.82 and 0.88 were calculated on the Matrices for the grade seven, eight and nine students respectively. The Gates MacGinitie Reading Survey and the Wisconsin Contemporary Test of Elementary Mathematics were used as criteria in calculating respective concurrent validity measures of 0.45 and 0.59. Similarly construct validity measures of 0.34 and 0.61 were calculated using Canadian Lorge-Thorndike verbal and nonverbal intelligence scores as the criteria. While all correlations calculated were significant, the correlation between the verbal intelligence scores and the Progressive Matrices is considered to be low.

Scores on the Progressive Matrices were essentially unaffected by socio-economic status. However, differences between the native and the non-native sample were found to be significant at the 0.01 level for the grade seven and eight students. The difference was not found to be significant at the grade nine level.

ACKNOWLEDGEMENTS

The writer wishes to acknowledge and thank [illegible] persons who contributed generously of themselves [illegible] possible the completion of this dissertation.

I express my thanks to my faculty advisor and dissertation chairman Dr. E.E. Fox and the other supervisory and examining members of my candidacy and final examination committees: Drs. G. Fitzsimmons, V. Nyberg, B. Schwartz, R.B. Alderman and the external examiner, J.E. Bicknell. I thank also Dr. T Maguire and C. Park for their invaluable assistance in the analysis of the data.

Appreciation is extended to H.A. MacNeil and the members of the Board plus the principals, teachers and students in the Edmonton Separate School System for their financial assistance, permission and co-operation which allowed me to carry out this research.

A deep appreciation is extended to my wife, Pat, for her understanding, co-operation and encouragement and for the many hours she devoted to the typing of the thesis. Finally appreciation to my children, Shannon, Kelly, Darci and Michael, who are a source of inspiration to their father.

It is to my wife and children that I wish to dedicate this document.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION TO THE STUDY

I. INTRODUCTION

In recent years there has been a growing awareness on the part of educators of the child who does not fit the stereotype of the average or normal student. Parent groups, governments, school boards, universities, and teachers have promoted the development of special education settings which would facilitate the optimum educational experience for each child whatever his handicaps or strengths. In reaction to this pressure school systems have established, in addition to a variety of "levels" within the regular school program, special classes for the mentally retarded, emotionally disturbed, culturally deprived, neurologically impaired, and mentally gifted.

Because of these special classes there has arisen an increased interest in measuring a student's scholastic achievement in relationship to his intellectual ability. It is in fact suggested that most education philosophies promote the concept of adapting instruction to the needs and abilities of the individual students (Rattan & MacArthur, 1968). It appears therefore that educators should be cognizant of a student's ability and be capable of comparing this with his actual academic progress. Marked differences between the two could be an indication of a "learning problem" or a "teaching problem". Cronbach (1960) further suggests that mental ability tests ought to help locate undeveloped potential that a unique educational program might bring out.

Related to the foregoing is the ever present interest in prediction. Countless research projects have been carried out which attempt to determine how effectively specific tests or specific criterion predict future behavior or future success. Once again, to use the educational setting as an example, it has often been the practice to direct a student into a particular program and indeed a particular vocation on the basis of the results of a specific testing program. This point is reinforced by Evens (1970) and by Vernon (1970) when they claim that assessment of intellectural potential of pupils has prognostic as well as diagnostic value. That is, future academic or general success is often predicted on the basis of intellectual potential as determined through the use of an intelligence test. MacArthur (1968-a) summarizes this line of thinking when he states:

> ... to the extent that we can implement a philosophy calling for teaching procedures and curricula so adapted as to

> *maximize realization of the present intellectual potential of individuals, and if we are attempting to predict to distant and general goals, theoretical consideration of constructs likely to be involved in predictor-goal relationships become useful, perhaps definitely necessary. Hence a need for criteria for the construct validity of measures of present intellectual potential.* (p. 45)

II. BACKGROUND TO THE PROBLEM

It would appear that the search for an acceptable definition and measure of intelligence has, for many years, been one of the major pre-occupations in the field of psychology. It can also be seen from the foregoing that school systems are faced with a continuous and growing demand for accurately assessing, among other characteristics, the intellectual and cognitive abilities of its students. Results of intellectual assessments are often used to determine a student's placement in a special program or special setting. Having accepted this mandate, the school systems thus must cope with the problem of selecting the most viable methods and instruments available which can accurately assess the level of intellectual functioning of the students.

The decision on what instrument(s) to utilize depends on a number of factors: reliability and validity of the instrument(s), cost of the program both in terms of qualified personnel and in terms of direct financial outlay for materials, logistics of administration and scoring, functional utility of the results, numbers of students that can be assessed in specific time periods, and availability of normative data. It is generally agreed that individual intelligence testing tends to be more valuable than group testing. This is particularly true in clinical diagnosis. However, standardized group tests are often, of necessity, chosen by school systems because of the above mentioned factors. There is pressure therefore to find group tests which could serve as general screening instruments that would meet some of the needs as expressed in the opening paragraphs of this chapter; i.e. assessment of present level of intellectual functioning and prediction of future level of intellectual functioning.

III. PROGRESSIVE MATRICES

One such instrument that can be utilized as a group test in addition to being used as an individual or self-administered test is the Standard Progressive

Matrices which was developed and normed by Raven in Great Britain. This is a nonverbal intelligence test which was designed to measure Spearman's general factor of intelligence. (This "g" factor will be discussed in the following chapter.) The test is reputed to provide a measure of the capacity to form comparisons, to reason by analogy and to develop a logical method of thinking regardless of previously acquired information (Westby, 1953).

In order to counteract specific application limitations with the original test, Raven has developed two variations. The Progressive Matrices (1947) A, AB, B (Colored Progressive Matrices) was developed for use with very young children and mentally subnormal or impaired individuals, while the Progressive Matrices (1947) Sets I and II was designed to discriminate among elite individuals in the top quarter of the population. This study, however, will concern itself only with the (revised) original instrument, Standard Progressive Matrices (1938). Henceforth in this study the test will be referred to as the Progressive Matrices or simply the Matrices.

IV. STATEMENT OF THE PROBLEM

The Progressive Matrices is not meant as a substitute for the Binet or Wechsler tests, nor for any verbal or nonverbal group test of mental ability, however it is considered by Burke (1958) to be almost as useful in that it shows inter-correlations with such tests perhaps as high as they show with each other. While the Progressive Matrices has been extensively used in England, its use in North America has been comparatively limited. It is suggested that this lack of use results from the fairly general notion that the test requires improvement in regard to validity and normative data (Burke, 1958). As stated by Blumenkrantz, Wilkin and Tuddenham (1968):

> *... the normative sample though very large, is not well described in Raven's manual, and information on the relationship between the Progressive Matrices and other measures is only slowly accumulating.* (p. 931)

It is felt that if these improvements could be brought about, particularly in the area of normative data, that the Progressive Matrices could become a useful research and educational tool in North America.

While a number of research projects have been carried out in recent years in Canada using the Progressive Matrices (MacArthur, 1962; MacArthur, 1965; MacArthur, 1968-b; Rattan and MacArthur, 1968; West and MacArthur, 1964) these have been generally restricted to the examination of intellectual functioning in native children. There is a marked lack of reported research on the use of the Progressive Matrices in an urban Canadian setting. It would further appear that the actual use of the Progressive Matrices in the Edmonton area, a Canadian city of approximately one half million people, is almost negligible. Wechsler (1960) would suggest that this general lack of use is somewhat unfortunate as the Progressive Matrices could be a welcomed addition to psychometrics because it constitutes a definite contribution to the field. Westby (1953) reinforces this thinking when he "warmly" recommends the use of the test to North Americans.

Although the test as mentioned above is easy to administer, has a great deal of intrinsic motivation, is economical, and can be administered to groups by competent teachers, the lack of local norms tends to militate against its use in the Edmonton area. It is felt that if local norms were available that the Progressive Matrices could become an integral part of the total testing program in a large urban school system where a general screening instrument is required. This is particularly important where the school system has a relatively large number of Indian and Metis pupils, a relatively large number of ethnic students from various European countries, and a relatively diverse socio-economic population. A "culture-reduced" test such as the Progressive Matrices given in conjunction with the regular academic testing program could help to distinguish between present intellectual potential and present level of performance (Hebb, 1958). It could provide teachers and counsellors with a less biased estimate of the potential of not only pupils who are the product of different cultures and different socio-economic backgrounds but also of pupils who come from a background that is considered to be "average" (Pedrini, 1972).

V. PURPOSE OF THE STUDY

Based on the foregoing, the purpose of this research project was to provide information which would allow for the effective use of the Progressive Matrices in the Edmonton area. The specific objectives of the study were:

1. To develop normative data for the Progressive Matrices for use with the grade seven, eight, and nine pupils in the Edmonton Catholic School System which has a total student population slightly in excess of 30,000.
2. To examine the internal reliability of the Progressive Matrices.
3. To validate the Progressive Matrices as given to the grade eight students with the Canadian

Lorge-Thorndike Intelligence Test which is administered to the grade eight students as part of their regular testing program.

4. To validate the Progressive Matrices as given to the grade seven students with their scores on the standardized grade six reading and arithmetic tests. The standardized achievement tests involved are the Gates MacGinitie Reading Survey and the Wisconsin Contemporary Test of Elementary Mathematics.

5. To note the relationship between the socio-economic status of the students as determined by the Blishen Index and the scores obtained on the Progressive Matrices.

6. To compare the performance of a sample of native students with the performance of a sample of non-native students on the Progressive Matrices.

CHAPTER II

THEORETICAL CONSIDERATIONS AND RELATED LITERATURE

The presentation of related literature is divided into three major sections, each having a specific relationship to the problem in question. The initial section of this chapter outlines the research which deals with the history of intelligence testing. The second section reviews the literature on the nature of intelligence. This section outlines the positions taken by a number of the dominant theorists in the field of intelligence testing and research. Following immediately from this review an attempt is made to formulate a definition of intelligence. The final section briefly reviews the relationship between socio-economic status and success on intelligence tests and academic achievement tests. The literature dealing specifically with the Progressive Matrices will be dealt with in Chapter III.

I. THE SEARCH FOR INTELLIGENCE

The history of intelligence testing can be said to begin around the turn of the century with the work of Alfred Binet. The school authorities of Paris requested Binet to develop a set of tests which could be used in the objective identification of mentally deficient school children in order that they might be provided with a more appropriate curriculum. Binet and his associate, Simon devised a number of tests of attention, memory, discrimination, etc. which could be performed by average individuals (McMahon, 1974). By rank ordering these tests from easiest to hardest they attempted to discover each child's level of mental development. Binet and Simon felt that it was possible to: determine in a scientific manner the mental level (intelligence) of a child, compare this level with the level of the other children and thus determine by how many years a child might be retarded. Operating on this belief Binet and Simon, during the initial revisions of their test, developed the procedure of relating mental age and chronological age (McMahon, 1974).

One of the major shortcomings in Binet's tests was his failure to control for differences in the language and cultural background of the testees (Lindgren, Byrne, and Petrinovich, 1961). In order to overcome some of these factors psychologists at Stanford University undertook the development of the Stanford-Binet Intelligence Scale. This scale followed the guidelines established by Binet. Terman, whose influence on North American concepts of intelligence is probably unparalleled, devised the first of these scales in 1916.

In collaboration with Merrill he devised the second one in 1937. This second scale which was more extensive and consisted of two forms, L and M, became the most extensively used device for the individual testing of children's intelligence to that time. The concept of the intelligence quotient was introduced along with the 1916 revision following a suggestion by Terman's contemporary, Stern. In 1960 a final revision of the Stanford-Binet was undertaken by Terman and Merrill. This revision incorporated the "best" items from the L and M forms and became known as the Stanford-Binet, Form L-M. This test, with a few minor revisions, is the instrument in use today.

Shortly after the publication of the initial Stanford-Binet the military authorities in the United States were faced with the problem of determining how hundreds of thousands of recruits could be classified for the purpose of determining military assignments. At the request of the United States Army a number of members of the American Psychological Association set out to develop a method by which large numbers of recruits could be quickly assessed. Based on the work of Otis they developed the Army Alpha, and the Army Beta tests. The former is a verbal test while the latter is a nonverbal test designed for use with persons unable to read English and for illiterates. During the second world war two similar tests were introduced: Army General Classification Test and the Armed Forces Qualification Test. These tests like their predecessors were intrinsically interesting while their administration and scoring were easily carried out. The use of these tests on large numbers of recruits demonstrated the practicality and the value of group intelligence testing and opened the way for many tests of similar design (Lindgren, et al, 1961).

Following the army's success with group testing, similar group intelligence tests were developed and introduced into the public schools during the 1920's with the hope of improving the efficiency of education. The group tests appeared to offer a method by which the students could be easily and quickly classified according to ability. The theoretical rationale behind such a classification program was to adapt curriculum and methodology to the level of the students. While the idea has not worked out as well in practice as it did in theory, Lindgren, et al (1961) feel that the use of intelligence tests did introduce teachers to the concept of individual differences. This in turn has resulted in intelligence testing becoming a standard procedure in most school jurisdictions.

One of the more recently developed group tests is the Lorge-Thorndike Intelligence Test which was developed in 1954 with a Canadian version being developed in 1967. The original standardization group, which consisted of 136,000 school children, was selected to provide proportional representation of the various socio-economic levels in the U.S. (Lorge, Thorndike, & Hagen, 1972). The test consists of verbal and nonverbal items and provides such information as the individual's intelligence quotient and his percentile rank. Canadian norms have recently been published for the Lorge-Thorndike which has resulted in extensive use being made of the test in Canadian schools.

While the group tests provided a general screening device for the purpose of selection and classification they were not always functional for diagnostic purposes. The Stanford-Binet did provide some clues that aided in the diagnosis of various types of emotional and intellectual dysfunctioning. However it provided only one score and did not appear to be particularly adaptable for use with adults. There was also an extremely heavy weighting given to verbal material which appeared unfair to individuals who were not verbally inclined but appeared intellectually competent.

In an attempt to overcome some of the criticisms of the Stanford-Binet, Wechsler developed the Wechsler-Bellevue Intelligence Scale in 1939. The primary function of this scale was to provide an intelligence test suitable for adults. In 1949 Wechsler introduced the Wechsler Intelligence Scale for Children which was designed as a downward extension of the Wechsler-Bellevue. Wechsler in 1955 redesigned the Wechsler-Bellevue into the Wechsler Adult Intelligence Scale. The Wechsler scales are considered to be a descendent of the Binet-Simon scales and consist of both a verbal and a performance (nonverbal) scale.

One of Wechsler's primary motives behind the development of his tests was to produce an instrument that could be utilized in diagnosis of various mental and emotional disorders. Research, however, has not supported Wechsler's original claims that various subscore combinations have diagnostic significance. Experienced clinicians, however, have been able to gain a good deal of diagnostic information from analyzing the manner in which a subject responds and by analyzing the "qulaity" of responses (Lindgren et al, 1961).

While Binet and Simon were developing the first intelligence tests in France, Spearman was exploring intelligence from a different aspect in England. He and his co-workers were attempting to identify and isolate the "factors" that make up intelligence. In attempting to discover the nature of intelligence Spearman pioneered the factor analytic movement of modern psychometric research. His research also led directly to the work of Raven who explored possible techniques whereby intelligence could be measured while reducing the effects of environmental influences.

II. THE NATURE OF INTELLIGENCE

Upon examination of the literature it becomes apparent that there are three major models by which one can come to understand the nature of intelligence. The first model originates with a group of psychologists who view intelligence as consisting of two factors. They take as their basis the work of Spearman. The second model stems from the work of a group of psychologists who view intelligence as consisting of a number of individual factors. Among the most noteworthy of these theorists, who are referred to as "group factor" or "multiple factor" theorists, are Thurstone and Guilford. Wechsler can also be considered as a member of this group although he does not take as rigid a position. The third group which includes Hebb, MacArthur and Vernon attempts to reconcile the two opposing schools.

A. *Two Factor Model of Intelligence*

The two factor model of intelligence originated with the work of Spearman in England around the turn of the century. Spearman (1927) viewed intelligence as consisting primarily of a general factor "g" which is common to every ability plus specific factors "s".

Each "s" factor is specific to a particular task or ability and can be considered as independent of "g" and of all other "s's". The two factor theory accounts for the fact that mental tests generally show positive intercorrelations with each other (Krech, Crutchfield, and Livson, 1970). In explaining his theory, Spearman posited that many different skills "tap" the common pool of general ability "g". Such diverse abilities as musical ability, mathematical ability, spelling ability, etc. show a correlation with each other (admittedly this is often low) because they all require a certain amount of the common ability "g" (Munn, 1966). But skills such as mathematics require, in addition to "g", specific mathematical abilities "s" which might include facility with numbers, ability to factor, ability to multiply, etc.

Terman (1954) working in America, supported Spearman's concept of intelligence when he stated:

> *. . . that to achieve greatly in almost any field, talents have to be backed up by a lot of Spearman's "g", by which is meant the kind of general intelligence that requires ability to form many sharply defined concepts, to manipulate them and to perceive subtle relationships between them.* (p. 224)

In like manner Cattell (1943) in an extensive review of the theory relating to the nature of intelligence defended Spearman's "g" as a useful construct both in understanding intelligence and in intelligence testing.

Figure 1 illustrates the basis for correlation among tests according to Spearman's theory (Anastasi, 1968). It can be seen that tests 1 and 2 correlate highly with each other since each is highly saturated with "g", as shown by the shaded areas. The white area in each test represents specific abilities plus error variance. Test 3 which contains very little "g" is not highly correlated with the other two tests.

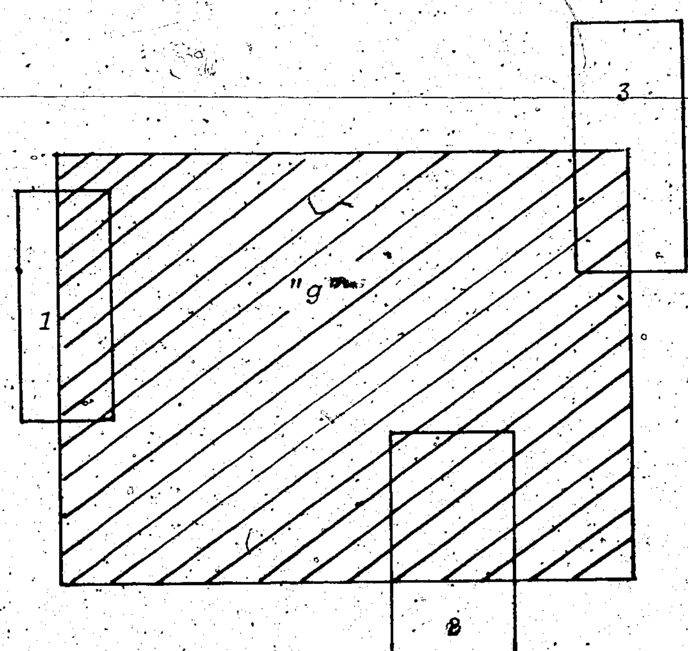


Figure 1

Two-Factor Model of Intelligence

Spearman's view of one general factor of intelligence has been challenged with the claim that there are other factors besides "g" which account for intercorrelations between tests of intelligence. However, his demonstration of the existence of at least one pervasive factor in all performance requiring intellectual ability is considered to be one of the great discoveries of psychology (Wechsler, 1958). In addition to this major discovery Spearman was the first psychologist to apply factor analysis in the study of intellectual abilities. Through factor analysis Spearman was able to apply a "stringent" test to his notion that there is one general ability present in all abilities. He developed and applied the test of tetrad differences. He felt that if there is a single ability common to four tests the relationship among the correlations would be:

$$r_{1,2}\, r_{3,4} - r_{1,3}\, r_{2,4} = 0$$

As explained by Deese (1967) the equation shows that the product of the correlation between tests 1 and 2 and the correlation between tests 3 and 4 is exactly equal to the product of the correlation between tests 1 and 3 and the correlation between tests 2 and 4. The equation is expressed such that the difference between the cross products of the correlations is zero.

The difference should be zero if all the tests have just one common factor. Deese (1967) points out however, that when a single general factor is removed, some correlation between the tests remain (except in very rare and unusual cases). While proponents of Spearman claim that this residual correlation is a function of statistical error it is generally accepted that there are other factors besides "g" that account for intercorrelations between tests of intelligence (Wechsler, 1958). These residual correlations led other psychologists to conclude that Spearman was only partially correct in regard to his general factor of intelligence. Spearman himself in his later formulations also included certain group factors in his theory of intelligence (Anastasi, 1968). These included such factors as arithmetic, mechanical, and linguistic abilities. This position however did not suggest a reduction in the importance or prevalence of "g".

B. *Multiple Factor Model of Intelligence*

In contrast to the two factor theorists a number of psychologists have taken exception to the claim that there is a general intelligence factor. They feel that what Spearman refers to as "g" is itself analyzable into a number of subsidiary factors. One of these psychologists, Thurstone, came to the conclusion that because we cannot account for all the correlation between tests by just one factor there must be two or more such factors.

Figure 2 illustrates the intercorrelations among five tests in terms of a multiple factor model. The correlations of tests 1,2, and 3 with each other results from their common loadings with the verbal factor (V). In like manner the correlation between tests 3 and 5 results from the spatial factor (S). The correlation between tests 4 and 5 results from the number factor (N). It can be seen that tests 3 and 5 are more complex in that they have a loading in more than one factor.

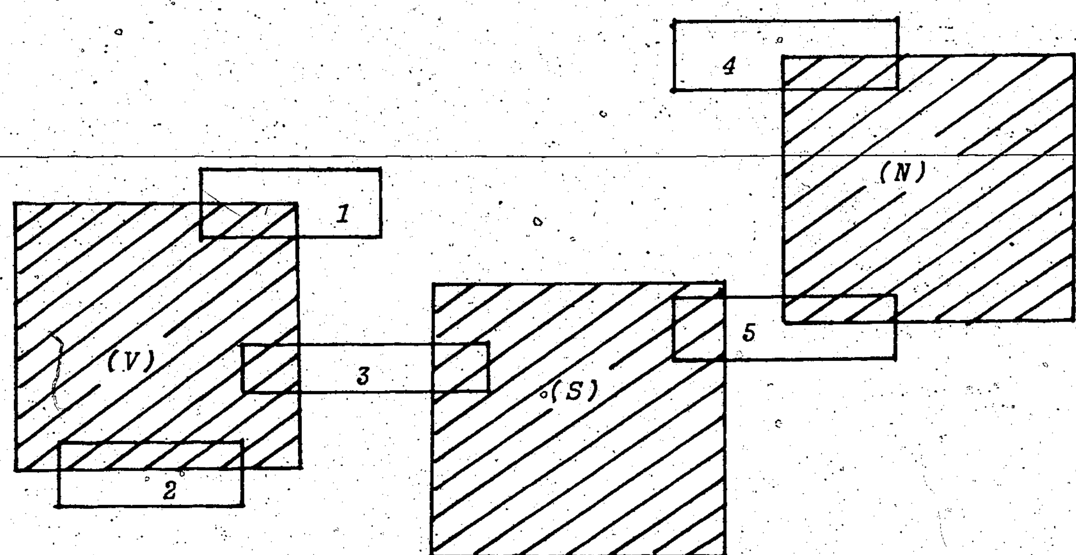


Figure 2

Multiple-factor Model of Intelligence

Spearman's experimentation with the tetrad-difference equations led Thurstone to a new method of factor analysis in an attempt to isolate the factors of intelligence. Thurstone (1947) developed a general procedure called multiple factor analysis for determining the pattern of factors that accounts for the correlations between different tests. In this procedure one common factor is first extracted from all the tests and then the residuals are computed from the correlation coefficients. A second factor is extracted from the residuals and so on until the residuals have been almost completely exhausted. Through this procedure a set of factors can be identified plus the weighting each test has in a particular factor.

Thurstone (1938) postulated, through his work with factor analysis, that there were seven primary mental abilities that could account for the mental capacity of man. These included the ability to perform the four fundamental arithmetic processes (number ability), the ability to write and speak with ease (word fluency), the understanding of ideas in word form (verbal meaning), the ability to make recall (memory), the ability to solve problems (reasoning), the ability to perceive size and spatial relationships correctly (spatial perceptions), and the ability to identify stimulus objects quickly (perceptual speed). The Thurstone Primary Mental Ability tests were developed to assess each of these mental abilities. Thurstone, however did not find these primary mental abilities to be independent of each other as a large residual variance remained after factor analysis. Wechsler (1943) claims that this might be due to the insufficient number of variables entering into the correlation matrix. The fact that these abilities were shown to correlate positively with each other tends to support Spearman's model of a common factor in each ability (Krech, Crutchfield, and Livson, 1970). This opinion reinforces the findings ~~of~~ Morris (1939) who after isolating thirty two variables was still able to account for only thirty-five percent of his total correlational variance.

Guilford (1959) also promotes a group factor theory of intelligence. His model is organized around three main aspects of human functioning: operations, content and products. An operation is a major intellectual process and includes such things as evaluation, convergent thinking, divergent thinking, memory and cognition. An operation is performed upon certain kinds of information called content which may be figural, symbolic, semantic, or behavioral. Applying an operation to content yields a product. The product is the form that information takes once it is processed. It may yield units, classes, relations systems, transformations or implications (Guilford, 1959). It can be seen that Guilford's model of intelligence involves 120 abilities; eighty of these have been identified through tests (Lefrancois, 1972). Krech et al (1970) suggests that the practical test of the validity of this proliferation of factors (and of tests to measure them), lies in their ability to predict behavior better than is possible by a single test score. Krech et al (1970) point out however that available evidence is not conclusive in this regard. In like manner they suggest that this group factor theory has not led to a better theoretical understanding of intelligence.

Wechsler (1958) not unlike Thurstone and Guilford views intelligence as consisting of separate identifiable facets which can be individually measured. In his earlier writings, Wechsler (1943) noted that attempts to appraise intelligence in terms of global capacity, that is, the ability or abilities to deal effectively with any specific situation, results in incomplete measures of the individual's capacity for intellectual behavior. He claims the reason for this incomplete or inaccurate measure of intellectual ability is the fact that intelligence tests, as now constituted, measure effectively only a portion of the capacities that enter into intelligent behavior. In other words, Wechsler feels that the limitations of intelligence tests are due to their deficiencies of content.

Wechsler differs somewhat from Thurstone and Guilford in that he does not deny the existence of a general factor in explaining the nature of intelligence. He feels however that the measure of this factor is best achieved through the assessment of specific skills such as verbal ability, abstract reasoning, arithmetical skills, etc. He also stresses that there are additional characteristics which also have a direct influence on intelligent behavior. He refers to these as non-intellective factors in general intelligence; they include all affective and conative abilities which in any way enter into global behavior. Stated differently, Wechsler (1958) sees intelligence as the aggregate or global capacity of an individual to act purposefully, rationally, and effectively. Intelligence is global because it is composed of elements or abilities which are qualitatively different but which at the same time are dependent and interrelated with each other. Wechsler (1958) feels that while intelligence is not a mere sum of intellectual abilities, the only way that it can be evaluated is by the measurement of the various abilities which make up intelligence. That is, intelligence cannot be evaluated as a unitary characteristic.

C. *Hierarchical Model of Intelligence*

A third school of thought presents a model of the nature of intelligence which attempts to incorporate the best features of the two factor model and the multiple factor model. Four of the major proponents of this model include Burt (1949), Hebb (1958), MacArthur (1968-b), and Vernon (1965). This model attempts to explain the nature of intelligence by combining general intellectual ability, which has a high index of heredability (Pezzula, 1972), and special intellectual abilities in a hierarchical structure.

At the top of the hierarchical model is general intellectual ability which is similiar to Spearman's "g". Further down the hierarchy are group factors similar to Guilford's concepts of operations, content and products, and to Thurstone's ability factors of number ability, word fluency, etc. The emphasis on these factors varies depending upon the tasks and persons defining them (MacArthur, 1968-b). Vernon (1965) points out that after removing the general factor from the top of the hierarchy (by some factor analytic procedure) the positive residual correlations always fall into two main categories: the verbal-educative group and the spatial-practical-mechanical group. Vernon suggests, and MacArthur agrees, that these second order (hierarchical) abilities can be further sub-divided into an almost unlimited number of increasingly specific abilities. The hierarchical model is diagramatically represented in figure 3.

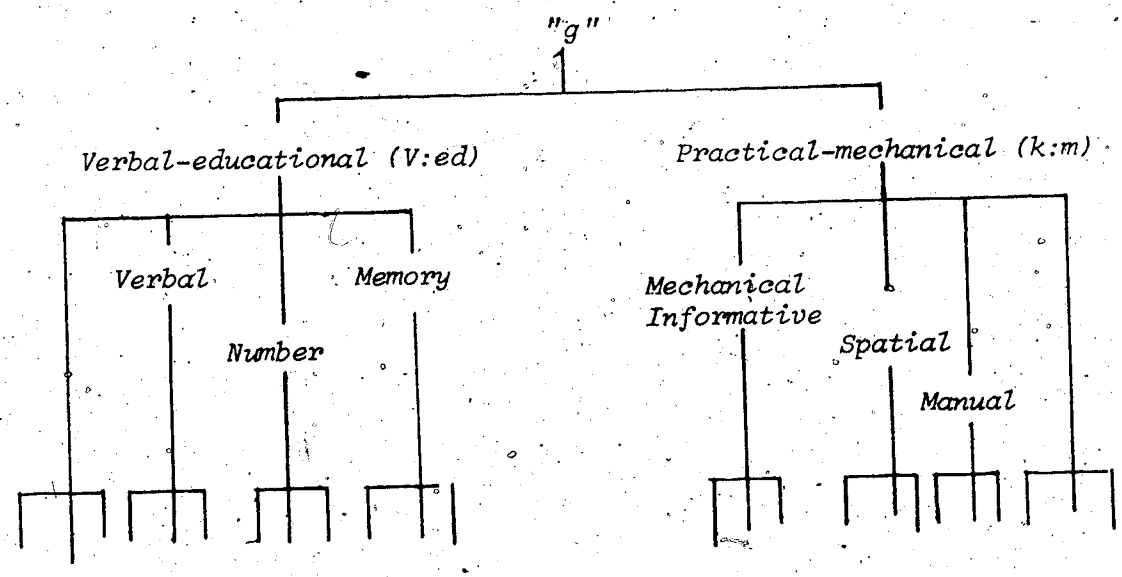


Figure 3

Hierarchical Model of Intelligence

Both Vernon (1965) and Krech et al (1970) feel that the issue between the two factor and multiple factor models is perhaps a non-issue when viewed in terms of the hierarchical model. They felt that both Thurstone and Guilford erred in two areas in their interpretation of intelligence. The first of these is concerned with the population that is usually dealt with in large scale intelligence testing research. These authors suggest that the reason for the popularity of the multiple factor approach, which in its pure form denies the existence of "g", is because much of the work in intellectual assessments is done with homogeneous groups such as "regular" school students, college students, armed forces personnel, etc. The range of "g" in these groups tends to be so restricted that in effect only group factors can be obtained through factor analysis. Vernon (1965) stresses this point when he claims that in most research, particularly in that involving older subjects, the sample is highly selected and less representative.

Vernon (1965) points out the second shortcoming in the interpretation of intelligence by the multiple factor theorists. He feels, and this is supported by Burt's (1940) writings, that Thurstone and Guilford erroneously regarded primary intellectual factors as fundamental components of the mind, which in combination make up all the important human capacities.

Spearman's position is also "corrected" by Vernon (1965). He felt that Spearman erred by viewing "g" as being determinate and unchanging other than that which results from physiological maturation. While "g" has a different loading on various intelligence tests it must not be viewed as a definite entity or autonomous mental faculty (Hebb, 1958). Hebb suggests that it be thought of in terms of a cumulative formation of more and more complex and flexible phase sequences or in Piagetian terms, schemata. That is, "g" grows, develops, and matures as a result of interaction between the individual and his environment. Thus "g" at any one point in time is the product of the innate potential for cognitive development and the cumulative influence of environmental experiences. It follows from this view of intelligence that nature and nurture are inseparable in the production of intelligent behavior (West and MacArthur, 1964). It is this developed and developing "g" that is tested with instruments such as the Progressive Matrices (Thorsen, 1971).

Vernon summarizes his position regarding the non-issue between two factor and multiple factor models by pointing out that a useful and accurate assessment of an individual's ability can be determined by applying the hierarchical model. He suggests that a test of "g" such as the Progressive Matrices combined with a test of verbal ability could provide an accurate assessment of present intellectual functioning. This could be supplemented by a specific group-factor test such as mechanical, clerical, etc. if a specific skill was being sought. In other words measures of ability that are high in the hierarchy generally have better external validity and more generalizability to living, whereas many of the primary factors are specific to particular tasks (Cronbach & Rajaratnam, 1963).

D. *Intelligence Defined*

In concluding this section on the nature of intelligence a theoretical definition of intelligence will be proposed. Historically the pragmatic nature of intelligence tests has overshadowed the theoretical need to adequately define intelligence. While psychologists have failed to come to terms with the problem of definition, the issue has resulted in extensive debate, much of which has been extremely useful, such as the nature-nurture issue, the two factor-multiple factor issue, etc.

The definition of intelligence that will be adopted for this study is derived primarily from the work of Hebb (1958) and MacArthur (1968-a). The definition consists of two parts, A and B. Intelligence A is the innate potential for cognitive development. This potential can be looked on as having two facets: A, the innate potential for cognitive development with which an individual is born and A^1, the present potential of an individual for future cognitive development, assuming that optimum future treatment is adapted to bring out that potential. Secondly, intelligence B is a general or average level of development of ability to perceive, to learn, to solve problems, to think, and to adapt. This second part of the definition, B, involves the personality factors that Wechsler (1958) feels are so important to intellectual functioning.

West and MacArthur (1964) suggest that while it is impossible to directly measure intelligence A or A[1], it is possible to measure intelligence B. If in doing so the variance in intelligence B which is attributable to environmental factors is minimized, the resulting score will be a reasonable measure of intelligence A[1] which is the present potential for future development, assuming optimum conditions. This position is reinforced by Rapaport, Gill, and Schafer (1968) who believe that when a subject takes an intelligence test, his performance represents his efficiency of functioning at that time. They go on to state that this efficiency or this intellectual potential is not necessarily a final and unchangeable characteristic. Like West and MacArthur (1964) they feel it will be greatly influenced by future environmental stimulation. West and MacArthur (1964) found that the Progressive Matrices was a test that, through reducing cultural variance, was able to provide a reasonable estimate of present intellectual potential.

III. SOCIO-ECONOMIC STATUS

It is a generally accepted hypothesis that intelligence is the product of both hereditary endowment and environmental influences. While the question of the relative importance of heredity and environment is far from settled it would appear that they both exert an essential influence on intelligence. Hebb (1958) stresses this point when he states:

> ... *to ask how much heredity contributes to intelligence is like asking how much the width of a field contributes to its area.* (p. 120)

Elley (1961) concludes that their importance varies to the extent that one is held constant while the other changes.

Socio-economic status is one such environmental variable that has received extensive attention in the research literature with the conclusion that there is a marked positive relationship between socio-economic status and intelligence test scores (Anastasi, 1958; E Elley, 1961; Hebb, 1958; Sperrozzo and Wilkins, 1959; Tesi and Montemagni, 1962; Tyler, 1953). Related studies have also shown similar relationships between socio-economic status and school achievement (Anderson, 1971); delinquency (Scarfe, 1974) and various other indices of "social adjustment". These studies were further confirmed by Koubekova and Miklova (1973) in an extensive study involving 20,000 students. They reported that parents' education and occupation were both significantly related to their children's intelligence. They also noted that parents' occupation was an important determinant of their children's academic success. Income per family, standard of housing and cultural setting were also found to be related to academic success.

A number of instruments have been devised to assess socio-economic status, one of which is the Blishen Index (1958). This index or scale was developed from data derived from the Canadian National Census. It is based on occupation, mean income, and number of years of schooling. Olsen and MacArthur (1962) have used it to determine socio-economic status of junior high school students in Edmonton. The Blishen Index is considered to be an objective measure of socio-economic status suitable for use in Canada (Olsen, 1962).

CHAPTER III

STANDARD PROGRESSIVE MATRICES

In 1958 Burke made what appears to be an exhaustive and critical review of the literature available on the three forms of the Progressive Matrices tests--the Standard Progressive Matrices, the Progressive Matrices (1947) Set I and II, and the Colored Progressive Matrices. Since this research project was planned to utilize only the Progressive Matrices this review will be limited to that literature which is relevant to that instrument. The procedure will be to report on Burke's summary and to update his review by examining the literature from 1958 until the present.

I. PROGRESSIVE MATRICES: DESCRIPTION

As viewed by Westby (1953) the Progressive Matrices is:

> ... *a nonverbal test of a person's capacity at the time of the test to apprehend figures presented for his observation (Spearman's Principle of Experience), see relations between them (Spearman's Education of Relations), and conceive the correlative figures completing the systems of relations presented (Spearman's Education of Correlates).* (p. 418)

It is further described by Westby (1953) as a test which is simple to administer, usually enthralling, and generally acceptable to the dull and clever, the young and old, and the friendly and defensive. The test is made up of logically designed patterns (matrices) which serve to measure what Raven (1938) refers to as "innate educative ability ", "educative intelligence", or simply "education".

Wechsler (1960) views the Progressive Matrices as an excellent test of general intelligence. He also sees it as an extremely interesting attempt at assessing intelligence in that it is the first attempt at measuring intelligence in terms of or through a single intellectual function--visual perception. This is a marked departure from the Binet tests and the Wechsler tests which attempt to assess intelligence through a variety of intellectual functions such as memory, knowledge, perception, reasoning, spatial relations, etc. This departure is emphasized by Westby (1953) who stated:

> *The concerted attack by Spearman, Burt and the British "g" school of factor analysis upon the empirically composed heterogeneous test as developed by Binet and his American followers se-*

> *cured its first indubitably spectacular success in Raven's Progressive Matrices.* (p. 418)

The Progressive Matrices consists of sixty designs or matrices divided into five sets (A B C D and E) of twelve problems each. Each problem or matrix consists of a network of logical relations between two dimensional, visual geometrical forms. Each matrix has a "gap" which must be filled by the testee who selects the correct choice from six or eight alternatives printed below the matrix. The relations within the matrix usually allow for more than one method for analysing the problem. The themes employed are: (a) continuous patterns, (b) analogies between pairs of figures, (c) progressive alterations of patterns, (d) permutations of figures, and (e) resolution of figures into constituent parts.

Abstract reasoning by analogy is described by Carlson (1973) as the most advanced form of reasoning in that it requires logical solutions of the type involved in the operational solution of class inclusion problems. Therefore the solution of Progressive Matrices is not made by configural symmetry or thematic representation (graphic processes) but through the operation of multiplicative classification.

Carlson (1973) concludes therefore that the test is a suitable measure of higher thought processes that derive from analytic and integrating operations.

The test is designed so as to produce five sets each progressively graded in difficulty both between and within sets. The first problem in each set is intended to be self-evident and is succeeded by eleven problems of increasing difficulty. The items and sets are of a sufficient range of complexity to discriminate in a short testing time a sample of the general population (Westby, 1953). A final characteristic of the test is that it is considered to be culturally reduced in that it tends to reduce the influence of formal academic training, language development, socio-economic background, and cultural "deprivation" (West and MacArthur, 1964; Meeker and Meeker, 1973).

II. PROGRESSIVE MATRICES: HISTORY AND USE

It was the work of Spearman and Burt with their concept of a general factor of intelligence that could be considered the originating stimulus for the development of the Progressive Matrices by Raven (Burke, 1958). The test quickly grew to be the most widely used of British intelligence tests in that it was administered to over three million members of the military during the second world war. Its use in

North America has been comparitively limited. However it is proving to be increasingly useful in racial and cross-cultural studies as a measure of general intellectual potential (Jensen, 1959; MacArthur, 1968-a; Sperrozzo and Wilkins, 1959). Tesi and Montemagi (1962) have also compared socio-economic status and intelligence using the Progressive Matrices. Burke (1958) reports it as having been extensively utilized in research with special groups such as: the deaf, the cerebral palsied, mental defectives, psychiatric patients, child guidance clinic patients, university students and apprentices.

III. PROGRESSIVE MATRICES: CONSTRUCT VALIDITY

Raven's (1940) original purpose for designing the Progressive Matrices was to develop a test of a person's "present capacity to form comparisons, reason by analogy, and develop a logical method of thinking regardless of previbusly acquired information". Other English psychologists notably Spearman (1946), Vernon (1947) and Vincent (1952) view the Progressive Matrices as perhaps the purest test for measuring "g". West and MacArthur (1964), Rattan and MacArthur (1968) and MacDonald and Netherton (1969) reinforce this opinion with their findings that the test while not being culture free can certainly be looked on as a culture-reduced test and thus an acceptable test of "g". Rimoldi (1948) found that the Progressive Matrices was culture reduced in that it did not depend to any large extent upon the education or the amount of information possessed by the subjects. Burke (1958) however does not feel that proof exists to show that the test is a pure measure of Spearman's construct of "g".

Burke (1958) notes that the highest correlation found between the Progressive Matrices and other intelligence tests was .86 with the Terman-Merrill. This was reported by Raven himself in 1939. Smith (1958) in a more recent study with exceptional children found a correlation of .74 between the Progressive Matrices and the Stanford-Binet. Smith who views this as a moderate to high correlation is reinforced by Sitkei and Michael (1966) who report a correlation of .65 between the Stanford-Binet and Progressive Matrices which they find to be significant.

The correlations between Progressive Matrices and the Wechsler tests are very similar to that which was reported between the Progressive Matrices and the Binet tests. Burke (1958) in his review cites reported correlations of .75 and .74 with full scale scores on the Wechsler tests. McLeod and Rubin (1962) found significant correlations between the WAIS and the Progressive Matrices of .58 on the verbal scale; .68 on the performance scale and .67 on the full scale score. In a later study Purl and Curtis (1971) report correlations between the Matrices and the WISC subscales of .45, .60, and .58.

Studies conducted by MacArthur (1968-b) suggest a positive relationship between the Progressive Matrices and certain other group tests of mental ability. These include the Safran Culture-Reduced Intelligence Test, Cattell test of "g" Scale 2, and some sub-tests of the Lorge-Thorndike nonverbal Intelligence Tests.

IV. PROGRESSIVE MATRICES: CONCURRENT AND PREDICTIVE VALIDITY

Up until 1958 the available correlation data indicate fairly good concurrent and predictive validity for the Progressive Matrices (Burke, 1958). In more recent studies West and MacArthur (1964) and MacArthur (1968-b) report that the Progressive Matrices shows a moderate relationship with concurrent school achievement. Burke (1958) also found abundant evidence of concurrent validity for the Progressive Matrices. He found that it was able to discriminate over a wide range of groups which differed in intellectual capacity as determined by a variety of criteria. Rattan (1966) in his studies of Canadian natives noted a moderate relationship between the Progressive Matrices and current school achievement. Giles (1964) also found the test to have a consistently high validity correlation with achievement criteria.

V. PROGRESSIVE MATRICES: RELIABILITY

While Burke concluded that the Progressive Matrices does not have a convincingly high coefficient of reliability, Foulds and Raven (1948) report figures of .93 for late teens while Raven (1940) claims a reliability of .88 for children. Since Burke came to his conclusion Rattan and MacArthur (1968) have found the test to be acceptably reliable over three and four year periods and West (1962) found it to be more stable than conventional tests over a five year period. These results reaffirmed Olson's (1962) observation that the Progressive Matrices tended to be fairly constant over a four year period. Elley (1961) claims that the instrument produces results that are at least as stable, if not more so than an intelligence quotient obtained from the California Test of Mental Maturity.

VI. PROGRESSIVE MATRICES: NORMS

Percentile norms are provided for the Progressive Matrices for each half-year interval between eight and fourteen years. No norms are available for the years between 14 and 20. This in itself limits its use with a large proportion of junior high school students. The available norms are based on British samples of 1,407 children, 1,680 military personnel (male) and 2,192 civilian adults. Use of the test in several European countries suggests that existing norms are adequate. Studies in a number of non-European cultures, however, have raised doubts about the suitability of existing norms for groups with backgrounds differing from the original normative sample (Anastasi, 1968; Green and Ewert, 1955).

A number of authors have cautioned against using the existing norms in a North American urban setting (Anastasi, 1968; Westby, 1953). Raven himself suggests that his sample of children might have been of lower average intelligence than would have been found in large urban settings (Green and Ewert, 1955). Burke (1958) and Magary (1967) also question the legitimacy of applying Raven's norms to different populations.

In a more general frame of reference Westby, (1953); MacArthur, (1965); Rattan and MacArthur, (1968) and Wechsler (1960) all point out the possible errors that might result in applying a set of norms developed on one population to a second, perhaps distinctly different, population. Anastasi (1968) reinforces this position when she points out the value of developing standardized norms for more narrowly defined populations.

Cronbach (1960) notes that norms quickly become obsolete due to an increasing level of education and on-going social change. He further points out that the most useful norms are those that permit the comparison of an individual with his companions and competitors. Cronbach (1960) also suggests that the only fair basis for comparing schools is to judge each school against schools with similar organization, similar curricula, and similar promotion policies. He feels that published norms are rarely based on meaningful population segments.

Westby (1953) suggests that perhaps the logical procedure for the use of the Raven, and indeed for any testing program, is to construct norms regional to the population for which the instrument is being utilized. He feels, and this is also implied by Raven, that there are pitfalls in drawing conclusions from norms which have been constructed other than on the specific population in question. MacArthur (1965) emphasizes this point when he explains that a testing instrument must have appropriate norms or be amenable to the production of norms. Rattan and MacArthur (1968) further suggest that because conditions vary so much from one area to another that there could be an equal variance in predictive validity on specific instruments. It can be seen from the foregoing that most assessment instruments, if they are to remain viable, demand on-going and long term research in the areas outlined.

VII. CONCLUSION

In conclusion, it appears that the bulk of the available literature suggests that the Progressive Matrices has acceptable validity and reliability as a test of present intellectual functioning. Its increasing utilization with Canadian native populations reaffirms it as a culture-reduced test which is easily and efficiently administered and scored. It is questionable, however, if the norms developed by Raven would be applicable to children in a North American urban setting.

CHAPTER IV

METHODOLOGY AND EXPERIMENTAL DESIGN

The primary purpose of this research project was to develop a table of norms on the Progressive Matrices that would be applicable for use with the junior high school students in the Edmonton Separate School System. Based on the comments of Raven (1940), Green and Ewert (1955) and Anastasi (1968) it was hypothesized that the students in the current study would consistently score higher than those in the Raven study. Hypothesis I stated operationally was:

H_1 *The two distributions did not arise, through random sampling, from the same population.*

Hypothesis I stated in the null form was:

H_0 *The two distributions arose by random sampling from the same population.*

A related but secondary objective was to conduct two validation studies. The first of these was between the results obtained by the grade seven students on the Progressive Matrices and the results they obtained six months previously on standardized reading and arithmetic achievement tests. The second validation study was between the results obtained by the grade eight students on the Progressive Matrices and on the Canadian Lorge-Thorndike Intelligence Test. The two tests were written within three months of each other.

A third objective of the study was to note the relationship between the results obtained on the Progressive Matrices and the socio-economic status of the students. Based on the foregoing literature review it was postulated that socio-economic status would have only a minimal effect on students' performance on the Progressive Matrices.

The fourth objective of the study was to examine the Progressive Matrices in a cross cultural setting. The purpose of this aspect of the project was to determine if there was a significant difference between the scores obtained on the Progressive Matrices by a sample of native students and the scores obtained by the non-native sample. It was assumed that because of the reputed culture-reduced aspect of the Progressive Matrices that there would not be a significant difference between the two groups.

I. SAMPLE

The subjects used in this study were selected essentially from the grade seven, eight and nine student population of a large, state-supported urban school system. The school system has a total enrollment of slightly in excess of thirty thousand pupils. The sample, although selected by age, included approximately ten percent of the students in each junior high school (grade seven, eight and nine) in each of the twenty-nine schools in the system which has junior high school students. The subjects in each school and in each grade were randomly selected from the alphabetically arranged grade lists using a standard table of random numbers (Lordahl, 1967).

A total of 764 students was involved in the development of the age norms while 736 students were involved in the development of the grade norms. A small number of twelve year old students in grade six and fourteen year old students in grade ten were also selected. A two level sampling procedure as suggested by Lord (1959) was utilized in the selection of the grade six and grade ten students.

The majority of the population of native students in the twenty-nine junior high schools visited made up the native sample used in the study. One hundred and six native students were involved.

II. INSTRUMENTS

A. *Raven Progressive Matrices*

The Progressive Matrices is a nonverbal, culture-reduced test which measures a person's capacity to form comparisons, to reason by analogy and to think logically regardless of previously acquired information (Westby, 1953). It is reputed to be a measure of general intellectual ability (Burke, 1958; MacArthur, 1965; Vernon, 1947; Vincent, 1952).

B. *Canadian Lorge-Thorndike Intelligence Test*

The Canadian Lorge-Thorndike Intelligence Test was the criterion with which the grade eight Progressive Matrices was validated. It consists of a series of tests of abstract intelligence which is designed to measure an individual's ability to work with ideas and to form or identify relationships among ideas. The test consists of a verbal and a nonverbal battery. The verbal battery is made up of five subtests: Vocabulary, Sentence Completion, Verbal Classification, Verbal Analogy and Arithmetic Reasoning. The nonverbal battery consists of Figure Analogies, Figure Classification, and Numerical Relationships. Wright, Thorndike and Hagen (1972), claim that the two batteries involve tasks that consist of the following types of intelligent behavior:

1. dealing with abstract concepts
2. interpretation and application of symbols
3. forming relationships between concepts and symbols
4. flexibility in the organization of concepts and symbols

Nyberg (1969) reports a reliability of 0.85 for both the verbal and nonverbal batteries. He also notes that the Canadian Lorge-Thorndike compares favorably with other tests in predicting academic achievement of Edmonton Separate School children. Churchill and Smith (1966) report a validity coefficient between the Stanford-Binet form L-M and the Lorge-Thorndike verbal and nonverbal batteries of 0.79 and 0.65 respectively. They also found correlations of 0.84 and 0.65 between the Lorge-Thorndike verbal and nonverbal sections and the sixth grade Iowa Test of Basic Skills. Neufeld (1973) reports significant correlations between the Wechsler Intelligence Scale for Children and the Lorge-Thorndike of 0.74. While she did not find much communality between any of the subtests she did conclude that either test could be used to obtain an "over-all" intelligence score.

C. *Standardized Achievement Tests*

The two standardized achievement tests which were utilized as the criteria in validating the grade seven Progressive Matrices were the Gates MacGinitie Reading Test - Survey D (Gates & MacGinitie, 1965) and the Wisconsin Contemporary Test of Elementary Mathematics (DeVault, Fennema, Neufeld & Smith, 1968). These two tests are given to all grade six students in the Edmonton Separate School System as part of their regular testing program. The mathematics test has reported measures of internal consistency ranging from 0.88 to 0.91 and a validity coefficient of 0.75. No evidence of validity measures is available for the reading test however alternate form and split-half reliability ranging from 0.78 to 0.89 are reported.

D. *Blishen Index*

Socio-economic status was determined through the use of the Blishen Index (1968). This instrument produces what is considered to be an acceptable measure of socio-economic status in Canada (Olson, 1962). Blishen (1958) reports a mean correlation of 0.85 between his Canadian index and similar scales standarized in Great Britain, New Zealand, Japan and Germany. A slightly higher correlation of 0.94 was noted with the National Opinions Research Centre Index developed by Hatt (1953).

The socio-economic status of a household is normally determined by recording the occupation of the father and comparing it to the index as developed by Blishen. Where there was no father in the home the mother's occupation was noted.

III. DATA COLLECTION

All of the principals of the twenty-nine schools were contacted and their co-operation obtained for the selection and testing of their junior high school students. The selected subjects were tested in groups of from nine to thirty students depending on the school size. The group method of test administration has been reported previously and found to be reliable with no significant differences between group and individual scores (Green and Ewert, 1955). Selected students who were absent during the initial testing were tested on a return visit to the school. Those students who had moved from the school were not replaced through further selection. A total of forty nine students who were selected were thus not included in the study.

Each student was seated at an individual desk and provided with an individual test booklet and a separate answer sheet (Appendix I). Standardized instructions were given as outlined in the Guide to the Progressive Matrices sets A B C D and E (Raven, 1938). A maximum time period of one hour was allocated for completion of the test.

IV. DATA ANALYSIS

All the data were analysed using the 360/67 Computer at the University of Alberta. The preliminary analysis of the Progressive Matrices results showed by six month intervals and by grade: the number of children tested, the range of scores, the mean score, the standard deviation and the variance. The raw scores obtained were converted to percentiles using computer program *SPSS (Statistical Package for the Social Sciences). The percentile norms for the raw scores were listed by age and also by school grade. These were compared to Raven's percentiles using a Kolmogorov-Smirnov Test of Goodness of Fit (Guilford, 1965) and the relationships noted. A third table of norms was developed by converting the age percentiles into deviation intelligence scores. Finally a measure of internal consistency was calculated using the Kuder-Richardson formula 20 (Ferguson, 1966). In the second part of the study the Progressive Matrices raw scores obtained by the grade eight students were compared with their results on the Canadian Lorge-Thorndike Test by computing a Pearson Product-moment correlation coefficient. In like manner, the raw scores obtained by the grade seven students on the Progressive Matrices were compared to their results on the Gates MacGinitie Reading Survey and on the Wisconsin Contemporary Test of Elementary Mathematics. A third correlational study was conducted comparing the raw scores obtained by all students on the Progressive Matrices and their socio-economic status as measured by the Blishen Index. All correlation coefficients were calculated using computer program DESTO 5.

The final aspect of the study examined the differences, if any, between the performance of the native and the non-native subjects on the Progressive Matrices. This was accomplished by calculating the level of significance of the difference between two means for independent samples as outlined in Ferguson (1966). A two-tailed test was employed.

CHAPTER V

RESULTS AND CONCLUSIONS

I. NORMS

The research reported here deals with four salient aspects of the Raven Standard Progressive Matrices Intelligence Test: normative data of the matrices, validation of the matrices, socio-economic status and the matrices, and cross-cultural testing with the matrices. This chapter reports the results and conclusions of the research conducted while the following chapter deals with the implications of these results and the recommendations stemming from the results. As stated in chapter four the primary purpose of this project was to develop norms for the Progressive Matrices that would be applicable for use with the junior high school students, that is, those students in grades seven, eight and nine in the Edmonton Separate School System. These students cover a chronological range from approximately twelve to fifteen years in age. Three tables of norms have been developed:

Table I: *Grade norms for seven, eight and nine expressed in percentiles*

Table II: *Age norms for twelve, thirteen, and fourteen expressed in percentiles*

Table III: *Age norms for twelve, thirteen, and fourteen expressed as deviation intelligence scores with a mean of 100 and a standard deviation of 15.*

TABLE I

Percentile Norms For

Grades Seven, Eight and Nine

Raw Score	Grade 7	Grade 8	Grade 9
20			
21			
22			
23	1	-	-
24	1	-	-
25	2	-	-
26	3	-	-
27	3	-	-
28	3	1	1
29	4	2	1
30	4	2	1
31	7	2	1
32	9	2	1
33	10	4	2
34	12	4	4
35	13	5	4
36	16	7	6
37	19	8	6
38	20	9	7
39	24	11	10
40	27	15	11
41	32	20	15
42	38	24	17
43	44	31	23
44	50	37	28
45	60	41	33
46	64	49	38
47	70	53	45
48	76	60	52
49	82	67	60
50	86	75	68
51	90	80	73
52	95	86	79
53	97	90	84
54	98	95	90
55	99	99	95
56	99	99	98
57	-	-	99
58	-	-	99
59	-	-	-
60	-	-	-

TABLE II

Percentile Norms For

Ages Twelve, Thirteen and Fourteen

Raw Score	Chronological Age					
	12-0 12-5	12-6 12-11	13-0 13-5	13-6 13-11	14-0 14-5	14-6 14-11
20						
21						
22						
23	1	-	-	-	-	-
24	1	-	-	-	-	-
25	1	-	-	1	-	3
26	2	1	-	1	-	3
27	3	2	1	1	-	3
28	3	2	1	2	-	4
29	3	2	1	3	-	4
30	4	2	2	3	1	4
31	7	3	2	3	1	4
32	8	6	3	3	2	5
33	10	8	3	4	2	6
34	12	8	4	6	4	8
35	14	9	5	6	5	8
36	15	11	9	7	6	9
37	18	12	11	8	8	10
38	20	14	13	10	8	10
39	24	16	15	11	14	11
40	26	21	21	13	18	12
41	31	26	24	15	25	13
42	37	32	28	20	28	14
43	42	38	37	25	34	22
44	47	45	44	30	38	26
45	58	52	48	36	41	33
46	61	57	55	44	43	41
47	70	62	59	48	48	47
48	77	70	63	55	54	55
49	82	79	67	64	62	62
50	86	82	77	71	69	69
51	90	87	84	73	74	74
52	96	91	87	81	80	80
53	97	94	92	86	84	84
54	99	98	97	91	90	90
55	99	99	99	98	94	94
56	99	99	99	99	98	95
57	-	-	-	-	98	98
58	-	-	-	-	99	99
59	-	-	-	-	-	-
60	-	-	-	-	-	-

TABLE III

Deviation Intelligence Scores for Ages Twelve, Thirteen and Fourteen

Raw Score	Chronological Age					
	12-0 12-5	12-6 12-11	13-0 13-5	13-6 13-11	14-0 14-5	14-6 14-11
20	-	-	-	-	-	-
21	-	-	-	-	-	-
22	-	-	-	-	-	-
23	66	-	-	-	-	-
24	66	-	-	-	-	-
25	66	-	-	-	-	-
26	69	66	-	-	-	-
27	72	69	66	66	-	-
28	72	69	66	69	-	-
29	72	69	66	72	-	-
30	74	69	69	72	66	-
31	78	72	69	72	66	-
32	79	77	72	72	69	-
33	81	79	72	74	69	-
34	83	69	74	77	74	-
35	84	80	75	77	75	-
36	84	82	80	78	77	-
37	86	83	82	79	79	81
38	87	84	83	81	79	81
39	89	85	[illegible]	82	84	82
40	90	88	88	83	86	83
41	97	90	89	84	90	83
42	97	92	91	87	91	84
43	99	95	97	90	[illegible]	84
44	103	98	98	92	95	90
45	104	101	99	95	97	93
46	106	103	102	98	97	97
47	108	105	104	99	99	99
48	111	108	105	102	101	102
49	114	112	107	105	105	105
50	116	[illegible]	111	108	108	108
51	119	117	115	109	110	110
52	126	120	117	113	113	113
53	128	123	121	116	115	115
54	135	131	128	120	119	119
55	135	135	135	131	123	123
56	135	135	135	135	131	125
57	-	-	-	-	131	131
58	-	-	-	-	135	135
59	-	-	-	-	135	135
60	-	-	-	-	-	-

Measures of variability were calculated for the grade scores and for the age scores. The number of subjects, mean, range, standard deviation and variances of the grade scores are recorded in Table IV.

A measure of internal consistency was also calculated for the Progressive Matrices at each grade level using a Kuder-Richardson formula 20 (Ferguson, 1966). A K-R 20 reliability coefficient of 0.86 was obtained at the grade seven level; 0.82 at the grade eight level; and 0.88 at the grade nine level. Based on guidelines outlined by Guilford (1965) these figures which are also shown in Table IV suggest that the Progressive Matrices has acceptable reliability. These figures also compare favorably to those reported earlier in this study by Elley (1961), Foulds and Raven (1948) and Rattan and MacArthur (1968).

TABLE IV

Grade Norms Variability

Grade	N	Mean	Range	S.D.	Variance	K-R 20
7	249	43.33	44	7.12	50.71	0.86
8	248	46.23	38	6.01	36.09	0.82
9	239	47.34	46	6.58	43.26	0.88

In like manner the number of subjects, mean, range, standard deviation and variance of the age scores are summarized in Table V.

TABLE V

Age Norms Variability

Age	N	Mean	Range	S.D.	Variance
12-0 } 12-5	137	43.59	34	6.72	45.17
12-6 } 12-11	108	44.60	44	6.72	45.20
13-0 } 13-5	119	45.53	28	5.95	35.44
13-6 } 13-11	123	46.81	38	6.53	42.62
14-0 } 14-5	133	46.65	45	6.75	45.48
14-6 } 14-11	144	46.81	46	7.68	58.91

A comparison was made between the age norms reported by Raven (1938) and the norms developed in this study. This comparison is charted in Figures IV, V, and VI for ages twelve, thirteen and fourteen respectively. Examination of the figures shows that Raven's norms essentially parallel the norms developed in this project. Use of the Raven norms however, consistently results in a higher percentile placement. This difference ranges as high as 26.8 percentile points.

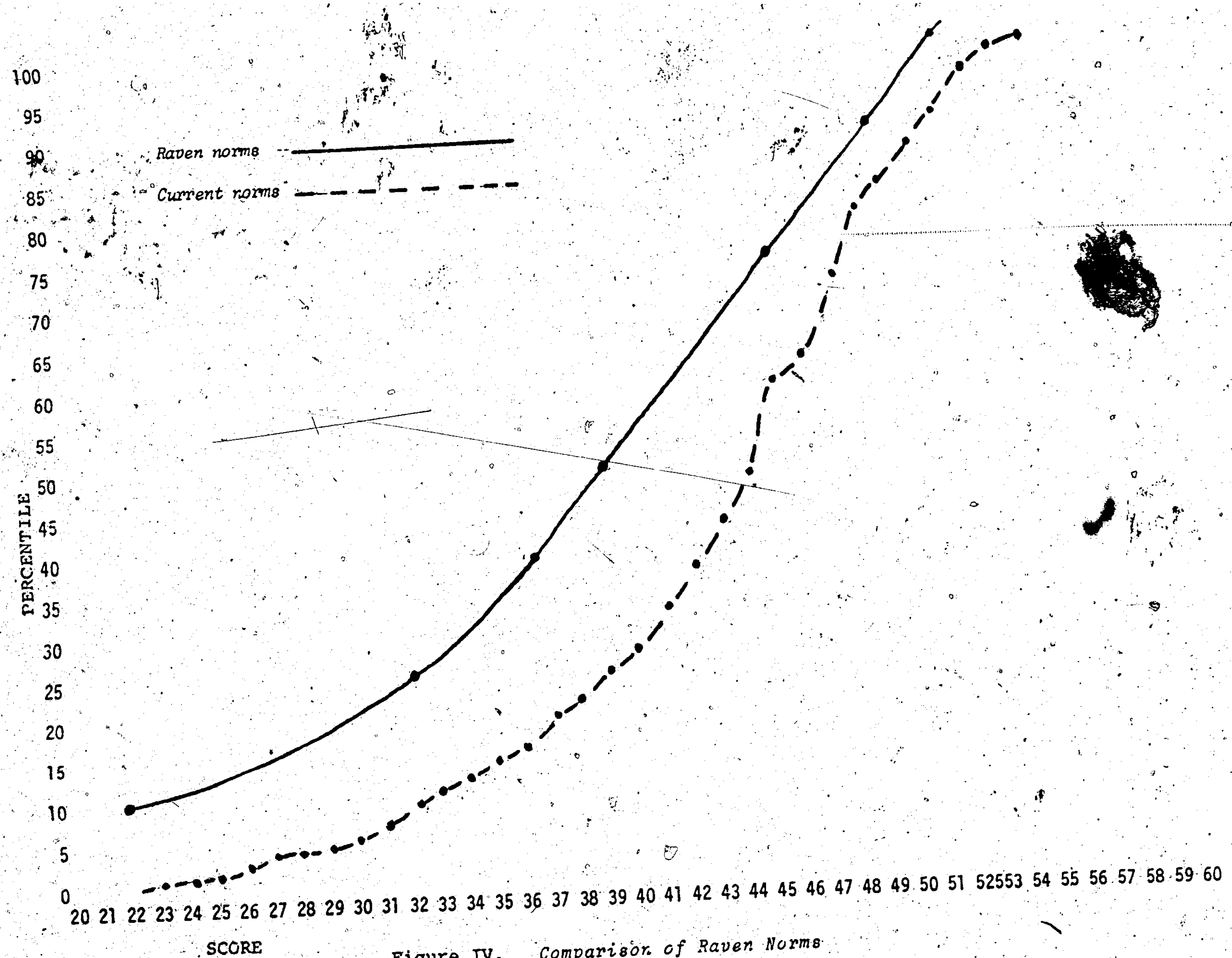


Figure IV. *Comparison of Raven Norms With Current Norms: Age 12*

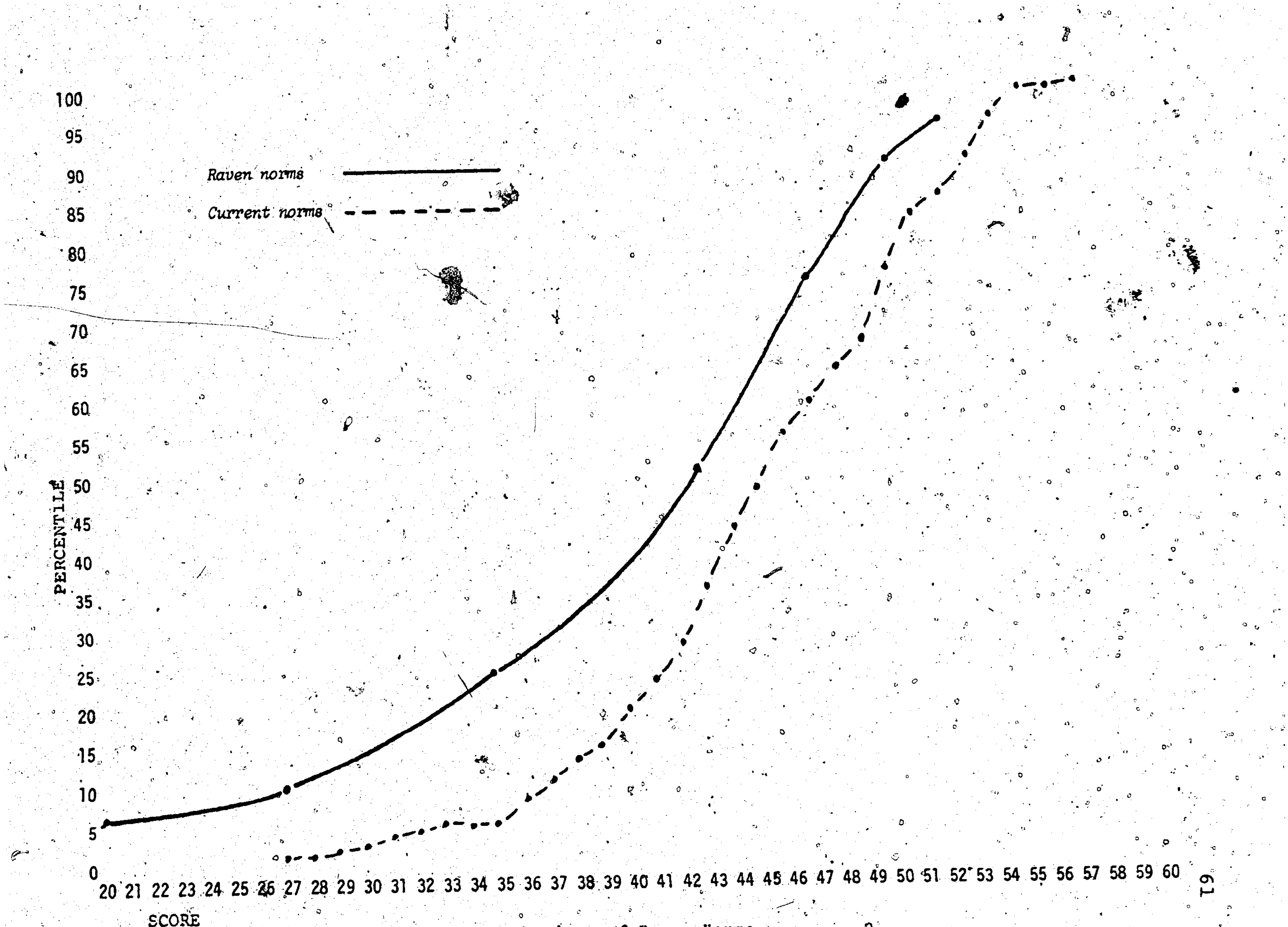


Figure V. *Comparison of Raven Norms With Current Norms: Age 13*

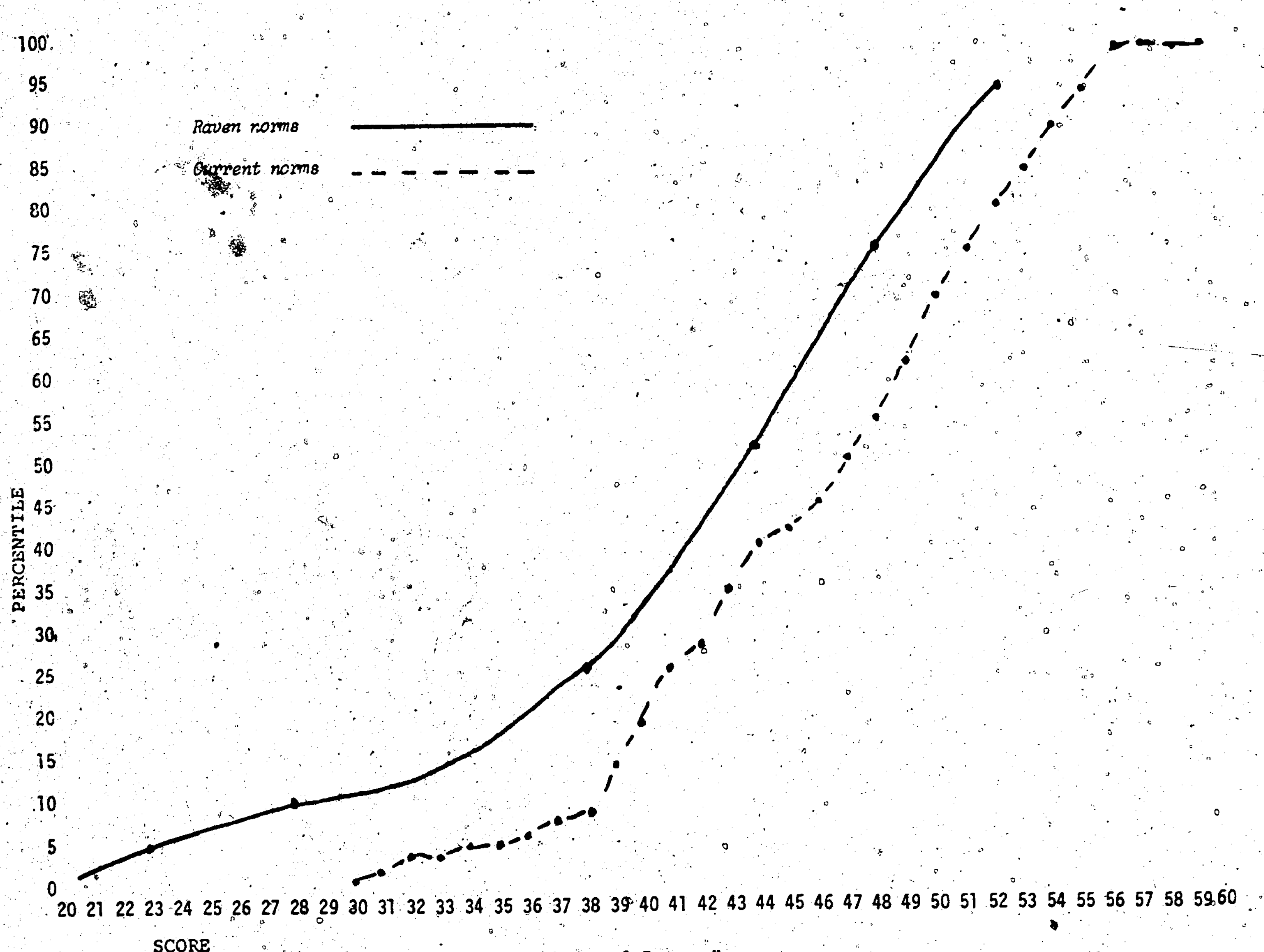


Figure VI. *Comparison of Raven Norms*

With Current Norms: Age 14

In order to assess the degree of difference between the two frequency distributions the Kolmogorov-Smirnov Test of Goodness of Fit was applied (Guilford, 1965). Using a one-tailed test it was found that the differences between the two distributions were significant at the 0.001 level for the grade seven and eight students and significant at the 0.01 level for the grade nine students. These results which are summarized in Table VI necessitated acceptance of Hypothesis I:

H_1 *The two distributions did not arise, through random sampling, from the same population.*

Conversely, the null hypothesis was rejected.

TABLE VI

Kolmogorov-Smirnov Test of Goodness of Fit Between Current Norms and Raven Norms

Current norms / Raven Norms	Grade 7	Grade 8	Grade 9
Grade 7	33.44*		
Grade 8		17.55*	
Grade 9			13.61**

* *Sig. at* $p \leq 0.001$

** *Sig. at* $p \leq 0.01$

II. VALIDATION OF THE MATRICES

In the second part of this study measures of concurrent and construct validity were calculated for the Progressive Matrices. The Gates MacGinitie Reading Test-Survey D and the Wisconsin Contemporary Test of Elementary Mathematics were used as the criteria for assessing concurrent validity. A Pearson product-moment correlation coefficient of r=0.45 was noted between the Matrices and the reading test. In like manner an r=0.59 was recorded between the Matrices and the mathematics test. Both correlations are considered significant.

These results show that the Progressive Matrices has acceptable concurrent validity (Guilford, 1965). This conclusion is consistent with those reported by Burke (1958), MacArthur (1968-b) and West and MacArthur (1964). The higher correlation noted between the Progressive Matrices and the mathematics scores can be explained by the highly analytic nature of the tasks involved. (Carlson, 1973; Rimoldi, 1948). This is opposed to the essentially verbal stimuli tasks that are involved in reading. A similar inference was made by Blumenkrantz et al (1968) when they stated:

> *It is to be expected that P.M., a nonverbal measure, should correlate somewhat lower with Reading and Vocabulary.* (p. 933)

The criteria used in calculating construct validity measures were the verbal and nonverbal batteries of the Canadian Lorge-Thorndike Intelligence Scale. Using the procedure noted above an r=0.34 was found between the Matrices and the verbal scale while an r=0.61 was found between the Matrices and the nonverbal scale.

The correlation coefficient of 0.61 between the Progressive Matrices and the nonverbal battery is similar to the correlations reported by McLeod and Rubin (1962) between the Matrices and the performance scale of the WAIS. The same authors however report markedly higher correlation between the Matrices and the verbal scale of the Wechsler test than the 0.34 found between the Mátrices and the verbal test used in this study. This lower correlation can be explained by the major emphasis in the Lorge-Thorndike on past, formal, verbal learning i.e. cultural and environmental factors involved in formal education. It is suggested that this is much more pronounced in the Lorge-Thorndike than in the Wechsler Tests. This results therefore in a lower correlation between a culture-loaded test such as the Lorge-Thorndike and a culture reduced test such as the Progressive Matrices.

Notwithstanding the foregoing the reported correlations between the Progres[illegible] Matrices and the Lorge-Thorndike, particularly t[illegible]al battery, lends credence to the construct validity of the Matrices. This conclusion is in essential agreement with the conclusion drawn in chapter three regarding the test's construct validity. The correlations calculated are summarized in Table VII.

TABLE VII

Pearson Product-Moment Correlations Between the Progressive Matrices And the Validation Criteria

	Reading	Mathematics	CL-T Verbal	CL-T N.Verbal
P.M.	0.45	0.59	0.34	0.61

III. SOCIO-ECONOMIC STATUS AND THE MATRICES

The third section of this project examined the relationship between the Progressive Matrices and the socio-economic status of the students by calculating a Pearson product-moment correlation coefficient. A significant correlation was found to exist between the Progressive Matrices and socio-economic status for the grade seven and eight students but not for the grade nine students. These results are summarized in Table VIII.

TABLE VIII

Pearson Product-Moment correlations

Between

The Progressive Matrices and the Blishen Index

	P.M. Grd. 7	P.M. Grd. 8	P.M. Grd.9
Blishen	0.16	0.30	0.00

While significant correlations were noted between the Progressive Matrices and the socio-economic status for both the grade seven and eight students these are not considered to be substantive. That is, the figures of r=0.16 and r=0.30, particularly the former, are small enough to be considered unimportant. Because of the comparatively large numbers involved in the study it would have been highly unlikely if a significant correlation did not exist. For practical purposes therefore it can be assumed that performance on the Progressive Matrices is not appreciably affected by socio-economic status.

The above conclusion is reinforced by the non-significant correlation calculated between the socio-economic status of the grade nine students and their scores on the Progressive Matrices. The foregoing observations lend further support to the reported literature which suggests that the Matrices test is in fact one which is minimally affected by cultural and environmental influences.

IV. CROSS CULTURAL TESTING AND THE MATRICES

The final objective of this project was to examine the Progressive Matrices in a cross-cultural setting to determine if a significant difference existed between the performance of a native (Canadian Indian) and a non-native sample of students. Using a t-test to determine the significance of difference between two means for independent samples (Ferguson, 1959) it was found that a significant difference did exist between the two populations at the grade seven and eight levels. No significant difference existed at the grade nine level. These findings are summarized in Table IX.

TABLE IX

Significance of Differences

Between Two Means

For Independent Samples

	Gr. 7 Non-native	Gr. 8 Non-native	Gr. 9 Non-native
Gr. 7 native	3.25*		
Gr. 8 native		2.92*	
Gr. 9 native			0.82

* *Sig. at* $p \leq .01$

An F test was also conducted to determine if there was a significant difference between the variances of the independent samples i.e. the native and the non-native samples (Ferguson, 1966). As noted in Table X no significant differences were found at the 10% level of confidence which lends further credence to the significant levels noted in Table IX.

TABLE X

Significance of Differences Between Variances of Independent Samples

	Gr. 7 Non-Native	Gr. 8 Non-Native	Gr. 9 Non-Native
Gr. 7 Native	1.28 (F=2.92)*		
Gr. 8 Native		1.55 (F=3.04)*	
Gr. 9 Native			1.71 (F=3.14)*

* Critical value of F

The above findings are contrary to the assumption made in chapter IV i.e. that there would not be a significant difference between the native and the non-native samples. A number of tentative explanations or assumptions are put forth in the following chapter which might account for these findings.

CHAPTER VI

IMPLICATIONS AND RECOMMENDATIONS

I. IMPLICATIONS

A. *Norms*

Based on the tripartite thinking that:

(a) the available norms on the Progressive Matrices are relatively incomplete

(b) test norms tend towards early obsolescence

(c) there is an inherent danger in applying a set of norms developed on one population to a second, perhaps distinctly different population

new norms were developed on the Progressive Matrices for use with the grade seven, eight and nine students in the Edmonton Separate School System. The findings show that the norms developed in the current study differ significantly from those developed by Raven. The current norms are significantly higher than the English norms.

This research also lends credence to the belief stated in chapter three that norms, if they are to be meaningful, should be developed on the population to which they are to be applied. These findings further justify the need for this and similar types of norming projects.

B. *Validation*

The significant correlation coefficients obtained between the Progressive Matrices and both the standardized tests of mathematics and reading show, using Guilford's (1965) rationale, that the Matrices has acceptable concurrent validity. This observation corresponds to one made by Kerr (1972) in which he noted that the Progressive Matrices shows promise as a discriminator of academic achievement. Relating the above to Cronbach's (1960) suggestion that concurrent and predictive validity are essentially the same concepts it is concluded that the probability is high that the Progressive Matrices has predictive validity in the areas of reading and mathematics.

The above would appear to be particularly true if the Progressive Matrices was utilized in conjunction with a vocabularly test as was originally suggested by Foulds and Raven (1948) and by Raven (1938). This would perhaps compensate for the tendency of the Matrices to correlate lower with verbal test batteries than with non-verbal test batteries as was noted both in this study and in the study reported by McLeod and Rubin (1962).

C. *Cross Cultural Testing*

Contrary to the assumption made in chapter four, significant differences were found between the performance of the native and the non-native students at the grade seven and eight levels. A number of tentative explanations can be put forth which might account for these findings. Initially it can be assumed that the nonsignificant difference at the grade nine level is due to the tendency of the native students to leave school. It is generally accepted that it is the "brighter" students who remain in school. This then would tend to reduce or counteract the effects of the different environmental and cultural backgrounds of the two groups which existed prior to grade nine.

A second assumption put forth to explain the significant differences is that the two groups come from markedly different environmental backgrounds and that the Progressive Matrices was unable to effectively reduce the effects of such influences. Such an assumption however, is not in keeping with the reported literature nor with the observations reported above which showed that the Progressive Matrices was essentially not affected by socio-economic status.

A third assumption, which is closely connected to the second, deals with the motivational and attitudinal factors that the native students bring to bear or perhaps more accurately, fail to bring to bear, on the writing of a test. If the native students are not motivated to do well on the test or if they have a highly impulsive response style they will score much below the highly motivated-reflective student (Kagan, 1971). A culture-reduced test however should take these factors, at least to some degree, into account.

A final and highly controversial assumption stems from the work of Jensen (1959). If one accepts Jensen's theories and if one accepts the meaning of intelligence as is traditionally used in the white European-American context it might be hypothesized that there are genetically based differences in intelligence and/or in methods of information processing between the native and non-native samples.

II. RECOMMENDATIONS

The above assumptions point out a number of shortcomings in the present study. These shortcomings in turn generate a number of suggestions for future research. Initially it appears necessary to determine to what extent the Progressive Matrices is culture-reduced. By comparing the results of a series of intelligence tests given concurrently to a native population it would be possible to calculate the extent to which the Progressive Matrices reduces the effect of cultural influences as compared to other tests. In this regard also it seems vital to compare the performance of a number of other sub-cultures on the Progressive Matrices. If there is a significant difference between each sub-culture and the total population it can be assumed that the Progressive Matrices is not a culture-reduced test.

In regard to the third assumption it would also appear important to examine the response style of the native subjects. The Progressive Matrices requires a highly reflective approach if one is to achieve a consistent degree of accuracy--impulsivity tends to lead to inaccuracy (Kagan, 1971). The third assumption also suggests that native students might require special or unique teaching programs. Guinagh (1969) points out that native students must be taught different types of learning skills than the non-native students if they are to achieve academically. It is important to determine therefore the skills in which these students are deficient. Future research in this area also should examine the relative importance of the three major factors in learning: previous knowledge, motivation and intelligence (Husen, 1967). If the significant differences between the native and the non-native groups are due to either a deficiency in previous knowledge or in motivation one would have to question if the Progressive Matrices is culture reduced. Such findings would also add strength to the above recommendation made by Guinagh (1969).

In addition to the foregoing it would prove worthwhile to carry out a series of "Jensen-type" experiments. This might help to determine if the native population is capable of "learning" to the same extent and with the same ease as are non-native students. Finally such research could perhaps offer educational suggestions for dealing with any noted differences in learning style.

One of the most obvious recommendations that arises is that the norms developed in this project should be utilized with the population in question. To use the norms developed by Raven would be most misleading and inaccurate. It follows from these findings that if the Progressive Matrices is to be effectively utilized local norms should be developed for each of the remaining grade and age levels. It is suggested that if this were to be accomplished that the Progressive Matrices could become a valuable asset to the local school systems. The value of the test is further emphasized when examined in light of its ease of administration, adaptability to group testing, relatively small capital outlay for materials, and its acceptable levels of reliability and validity.

The study further authenticates Anastasi's (1968) and Cronbach's (1960) position that norms are seldom developed on meaningfully specific populations. It is suggested therefore that where possible local norms should be developed for most testing instruments. If this is not possible great caution should be exercised in interpreting data from broadly based norms.

REFERENCES

REFERENCES

ANASTASI, A. Differential psychology. New York: MacMillan, 1958.

ANASTASI, A. Psychological testing. New York: MacMillan, 1968.

ANDERSON, B. Socio-economic status of students & school bureaucratization. J. of Educational Administration, 1971, 7, 12-23.

BLISHEN, B.R. The construction & use of an occupational class scale. Canadian J. of Economics & Political Science, 1958, 24, 519-531.

BLISHEN, B.R., JONES R.E., NAEGELE D., PORTER J. Canadian Society. Toronto: MacMillan, 1968.

BLUMENKRANTZ, J., WILKIN, W.R., TUDDENHAM, R.D. Relationship between the Progressive Matrices and A.G.C.T.-3A among older military personnel. Educational & Psychological Measurement, 1968, 28, 931-35.

BURKE, H.R. Raven Progressive Matrices: a review & critical evaluation. The J. of Genetic Psychology, 1958, 93, 199-228.

BURT, C. The factors of the mind. London: University of London Press, 1940.

BURT, C. The structure of the mind: a review of the results of factor analysis. British J. of Psychology, 1949, 19, 176-199.

CARLSON, J.S. A note on the relationship between Raven's Colored Progressive Matrices Test and operational thought. Psychology In The Schools, 1973, 10, 211-214.

CATTELL, R.B. The measurement of adult intelligence. Psychological Bulletin, 1943, 40, 153-193.

CRONBACH, L.J. Essentials of psychological testing. New York: Harper & Row, 1960.

CRONBACH, L.J., RAJARATNAM, N. & GLESER, G.C. Theory of generalizability: A liberalization of reliability theory. British J. of Statistical Psychology, 1963, 16, 137-163.

CHURCHILL, W.D., SMITH, S.E., The relationship between the 1960 Revised Stanford-Binet Intelligence Scale to intelligence & achievement test scores over a three year period. Educational & Psychological Measurement, 1966, 26, 1015-1020.

DEESE, J. General psychology. Boston: Allyn & Bacon, Inc., 1967.

DeVAULT, M., FENNEMA, E., NEUFELD, K., SMITH, L. Wisconsin Contemporary Test of Elementary Mathematics: Manual. Princeton: Personnel Press, 1968.

ELLEY, W.B. A comparative analysis of socio-economic bias in selected intelligence tests. Unpublished Ph.D. thesis, U. of Alberta, 1961.

EVENS, G.T. Intelligence, transfer and problem solving. In: On Intelligence. Edited by B. Dockrell. London: Methuen & Co., 1970.

FERGUSON, G. Statistical analysis in psychology and education. Toronto: McGraw-Hill, 1966.

FOULDS, G.A., RAVEN, J.C. Intellectual ability & occupational grade. Occupational Psychologist, 1948, 22, 197-203.

GATES, A., MacGINITIE, W. Gates-MacGinitie Reading Tests Teacher's Manual. New York: Teachers' College Press, 1965.

GILES, G. Predictive validity of Progressive Matrices and two other nonlanguage tests of mental ability. J. of Educational Measurement, 1964, 1, 65-7.

GREEN, M.W., & EWERT, J.C. Normative data on the Progressive Matrices (1947). J. of Consulting Psychology, 1955, 19, 139-142.

GUINAGH, J.P. An experimental study of basic learning ability and intelligence in low socio-economic populations. In: On Intelligence. Edited by B. Dockrell. London: Methuen & Co., 1970.

GUILFORD, J.P. Fundamental statistics in psychology and education. Toronto: McGraw-Hill, 1965.

GUILFORD, J.P. Three faces of intellect. American Psychologist, 1959, 14, 469-479.

HATT, P.K. Jobs & occupations: a popular evaluation. In: Class Status & Power. Edited by R. Bendix & S. M. Lipsed. Glencoe, Illinois, Free Press, 1953.

HEBB, D.O. A textbook of psychology. Philadelphia: Saunders, 1958.

HUSEN, T. International study of achievement in mathematics. In: On Intelligence. Edited by B. Dockrell. London: Methuen & Co., 1970.

JENSEN, A.R. A statistical note on racial differences in the progressive matrices. J. of Consulting Psychology, 1959, 23, 272.

KAGAN, J. Developmental studies in reflection & analysis. In: Perceptual Development in Children. Edited by Kidd & Revoire. Internal Universities Press, 1966.

KERR, A.S. Determinants of performance of the Bender Gestalt test and Raven's Progressive Matrices (1947) test. J. of Learning Disabilities, 1972, 5, 219-221.

KOUBEKOVA, E. & MIKLOVA, J. Socio-economic status and cultural determinants of mental abilities and school performance. Psychological Abstracts, 1973, 49, 5464.

KRECH, D., CRUTCHFIELD, R.S., LIVSON, N. Elements of psychology. New York: Alfred A. Knopf, 1970.

LEFRANCOIS, G.R. Psychology for teaching. Belmont, California: Wadsworth Co., 1972.

LINDGREN, H.C., BYRNE, D., PETRINOVICH, L. Psychology: an introduction to a behavioral science. New York: John Wiley & Sons, 1961.

LORD, F. Test norms and sampling theory. J. of Experimental Education, 1959, 27, 247.

LORDAHL, D. Modern statistics for behavioral sciences. New York: Ronald Press, 1967.

LORGE, I., THORNDIKE, R., HAGEN, E. Manual for administration: Canadian Lorge-Thorndike Intelligence Tests. Toronto: Thomas Nelson & Sons (Canada) Ltd., 1972.

MacARTHUR, R.S. Assessing the intellectual ability of Indian & Metis pupils at Fort Simpson. Ottawa: Department of Northern Affairs & National Resources, 1962.

MacARTHUR, R.S. Mackenzie norming project. Ottawa: Department of Northern Affairs & National Resources, 1965.

MacARTHUR, R.S. Some differential abilities of northern Canadian native youth. International J. of Psychology, 1968-a, 3, 43-50.

MacARTHUR, R.S. Assissing intellectual potential of native Canadian pupils: a summary. Alberta J. of Educational Research, 1968-b, 14, 115-122.

MAGARY, J.F. *School psychological services*. Englewood Cliffs: Prentice Hall, 1967.

McLEOD, H.N., RUBIN, J. Correlation between Raven P.M. & the WAIS. *J. of Consulting Psychology*, 1962, *26*, 190-91.

McMAHON, F. *Psychology, the hybrid science*. Englewood Cliffs, N.F.: Prentice-Hall, Inc. 1974.

MEEKER, M., MEEKER, R. Strategies for assessing intellectual patterns in Black, Anglo and Mexican-American boys--or any other children--and implications for education. Educational Resources Information Center, ED 084 909, 1973.

MORRIS, M.A. A critical analysis of certain performance tests. *J. of Genetic Psychology*, 1939, *54*, 35-105.

MUNN, N.L. Psychology. Boston: Houghton Mifflin Co., 1966.

NEUFELD, H. The relationship between the WISC & the Canadian Lorge-Thorndike Intelligence Tests. Unpublished M.Ed. thesis, U. of Alberta, 1973.

NYBERG, V.R. Canadian Lorge-Thorndike Intelligence Test--a user's appraisal. A paper presented at the Convention of the Canadian Guidance & Counsel-ling Association, Edmonton, Alberta, June, 1969.

OLSON, E.R. The influence of foreign language background on performance of selected intelligence tests. Unpublished M. Ed. thesis, U of Alberta, 1962.

OLSON, R.R. & MacARTHUR, R.S. The effect of foreign language background on intelligence test performance. Alberta J. of Educational Research, 1962, 8, 157-167.

PEDRINI, B. Intelligent intelligence testing. Educational Resources Information Center, ED 069 694, 1972.

PEZZULA, T. The heritability of Jensen's level I and level II and divergent thinking. Educational Resources Information Center, ED 061 542, 1972.

PURL, M., CURTIS, J. Raven Progressive Matrices test correl[illegible]s with measures of Lorge-Thorndike and [illegible]er intelligence tests. Educational Resources Information Center, ED 052 231, 1971.

RAPAPORT, D., GILL, M., SCHAFER, R. Diagnostic psychological testing. New York: International Universities Press, 1968.

RATTAN, M.S. Predictive validity & stability of measures OF intellectual potential for two samples of Indian-Metis & Eskimo children. Unpublished M. Ed. thesis. U of Alberta, 1966.

RATTAN, M.S., & MacARTHUR, R.S. Longitudinal prediction of school achievement for Metis & Eskimo pupils. Alberta J. of Educational Research, 1968, 14, 37-41.

RAVEN, J. <u>Guide to the Standard Progressive Matrices (1938)</u>, <u>Sets</u> <u>A</u> <u>B</u> <u>C</u> <u>D</u> <u>and</u> <u>E</u>. London: H.K. Lewis and Co. Ltd., 1938.

RAVEN, J. Matrix tests. <u>Mental Health</u>. 1940, <u>1</u>, 10-18.

RIMOLDI, H.J.A. A note on Raven's Progressive Matrices Test. <u>Educational & Psychological Measurement</u>. 1948, <u>8</u>, 347-52.

SCARFE, N. Can you teach hostile students? <u>Education Canada</u>. 1974, <u>2</u>, 20-25.

SITKEI, E.G., MICHAEL, W.B. Predictive relationships between items on the revised Stanford-Binet Intelligence Scale (SBIS), Form L-M, and total scores on Raven's Progressive Matrices (PM), between items on the PM and total scores on the SMIS, and between selected items on the two scales. <u>Educational and Psychological Measurement</u>. 1966, <u>26</u>, 501-508.

SMITH, L.M. The Progressive Matrices and the Stanford-Binet with exceptional children. <u>Exceptional Children</u>. 1958, <u>24</u>, 374-5.

SPEARMAN, C. <u>Abilities of Man</u>. New York: MacMillan, 1927.

SPEARMAN, C. Theory of a general factor. <u>British J. of Psychology</u>. 1946, <u>36</u>, 117-131, 1946.

SPERROZZO, G., WILKINS, W.L. Further normative data on the Progressive Matrices. [illegible] <u>Psychology</u>. 1958, <u>22</u>, 35-37.

SPERROZZO, G., WILKINS W.L. Racial differences on Progressive Matrices. J. of Consulting Psychology. 1959, 23, 273[illegible]4.

TERMAN, L.M. The discovery & enc[illegible]ent of exceptional talent. Americ[illegible]chologist. 1954 9, 221-230.

THORSEN, E. The heritability [illegible]G" and figural divergent thinking. Doctor's thesis, Boston College (Chestnut Hill, Mass.) (DAI 32:1351A), 1971.

THURSTONE, L.L. Primary mental abilities. Chicago: University of Chicago Press, 1938.

THURSTONE, L.L. Multiple factor analysis. Chicago: University of Chicago Press, 1947.

TESI, G., MONTEMAGI, G. Influence of town and country upon children's intelligence. British J. of Ed. Psych. 1962, 32, 151-8.

TYLER, F.T. Comments on the correlational analysis reported in intelligence & cultural differences. J. Educational Psychology. 1953, 44, 288-95.

VERNON, P.E. The variations of intelligence with occupation, age, & locality. British J. of Psychology, Statistical Section. 1947, 1, 52-63.

VERNON, P.E. Ability factors & environmental influences. American Psychologist. 1965, 20 723-33.

VERNON, P.E. Intelligence. In: On Intelligence. Edited by B. Dockrell. London: Methuen & Co., 1970.

VINCENT, D.V. The linear relationship between age & score of adults in intelligence tests. Occupational Psychology. 1952, 26, 243-249.

WECHSLER, D. Non-intellective factors in general intelligence. J. of Abnormal & Social Psychology. 1943, 38, 101-103.

WECHSLER, D. The measurement & appraisal of adult intelligence. Baltimore: The William & Wilkins Co., 1958.

WECHSLER, D. The third mental measurements yearbook. Ed: O.K. Buros, The Gryphon Press, Highland Park, New Jersey, 1960.

WEST, L.W. Assessing intellectual ability with a minimum of cultural biases for two samples of Metis & Indian children. Unpublished M. Ed. thesis, U. of Alberta, 1962.

WEST, L.W., & MacARTHUR, R.S. An evaluation of selected intelligence tests for two samples of Metis & Indian children. Alberta Journal of Educational Research. 1964, 10, 17-27.

WESTBY, G. The fourth mental measurements yearbook. Ed: O.K. Buros, The Gryphon Press, Highland Park, New Jersey, 1953.

WRIGHT, E.N., THORNDIKE, R.L., & HAGEN, E. *Technical supplement, Canadian Lorge-Thorndike Intelligence Tests*. Thomas Nelson & Sons (Canada) Ltd., 1972.

APPENDICES

APPENDIX I

NAME ______________________ Last ______ First ______ Middle

FACULTY or SCHOOL ______________________

AGE ______ Years YEAR or GRADE ______ ☐ Male ☐ Female DATE ______ Day Month Year

OCCUPATION ______________________

F ______________________

2. ______________________

M

Indicate response by placing a mark between the guidelines as shown in the example. Use HB pencil. Don't make marks longer than guidelines.

Example ▬

I. D. NUMBER

0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9

B

M

R

	0	1	2	3	4	5	6	7	8	9
G	0	1	2	3	4	5	6	7	8	9
V	0	1	2	3	4	5	6	7	8	9
	0	1	2	3	4	5	6	7	8	9
	0	1	2	3	4	5	6	7	8	9
NV	0	1	2	3	4	5	6	7	8	9
	0	1	2	3	4	5	6	7	8	9
	0	1	2	3	4	5	6	7	8	9

A

1.	0	1	2	3	4	5	6
2.	0	1	2	3	4	5	6
3.	0	1	2	3	4	5	6
4.	0	1	2	3	4	5	6
5.	0	1	2	3	4	5	6
6.	0	1	2	3	4	5	6
7.	0	1	2	3	4	5	6
8.	0	1	2	3	4	5	6
9.	0	1	2	3	4	5	6
10.	0	1	2	3	4	5	6
11.	0	1	2	3	4	5	6
12.	0	1	2	3	4	5	6

B

1.	0	1	2	3	4	5	6
2.	0	1	2	3	4	5	6
3.	0	1	2	3	4	5	6
4.	0	1	2	3	4	5	6
5.	0	1	2	3	4	5	6
6.	0	1	2	3	4	5	6
7.	0	1	2	3	4	5	6
8.	0	1	2	3	4	5	6
9.	0	1	2	3	4	5	6
10.	0	1	2	3	4	5	6
11.	0	1	2	3	4	5	6
12.	0	1	2	3	4	5	6

C

1.	0	1	2	3	4	5	6		
2.	0	1	2	3	4	5	6	7	8
3.	0	1	2	3	4	5	6	7	8
4.	0	1	2	3	4	5	6	7	8
5.	0	1	2	3	4	5	6	7	8
6.	0	1	2	3	4	5	6	7	8
7.	0	1	2	3	4	5	6	7	8
8.	0	1	2	3	4	5	6	7	8
9.	0	1	2	3	4	5	6	7	8
10.	0	1	2	3	4	5	6	7	8
11.	0	1	2	3	4	5	6	7	8
12.	0	1	2	3	4	5	6	7	8

D

1.	0	1	2	3	4	5	6	7	8
2.	0	1	2	3	4	5	6	7	8
3.	0	1	2	3	4	5	6	7	8
4.	0	1	2	3	4	5	6	7	8
5.	0	1	2	3	4	5	6	7	8
6.	0	1	2	3	4	5	6	7	8
7.	0	1	2	3	4	5	6	7	8
8.	0	1	2	3	4	5	6	7	8
9.	0	1	2	3	4	5	6	7	8
10.	0	1	2	3	4	5	6	7	8
11.	0	1	2	3	4	5	6	7	8
12.	0	1	2	3	4	5	6	7	8

E

1.	0	1	2	3	4	5	6	7	8
2.	0	1	2	3	4	5	6	7	8
3.	0	1	2	3	4	5	6	7	8
4.	0	1	2	3	4	5	6	7	8
5.	0	1	2	3	4	5	6	7	8
6.	0	1	2	3	4	5	6	7	8
7.	0	1	2	3	4	5	6	7	8
8.	0	1	2	3	4	5	6	7	8
9.	0	1	2	3	4	5	6	7	8
10.	0	1	2	3	4	5	6	7	8
11.	0	1	2	3	4	5	6	7	8
12.	0	1	2	3	4	5	6	7	8

CAUTION — ... THE BLACK TIMING LINES

OVERPRINT ANSWER SHEET

APPENDIX II

Grade Seven Frequency Distribution

Frequency															
56															
54										*					
52										*					
50										*					
48										*					
46										*					
44										*					
42										*					
40										*	*				
38										*	*				
36										*	*				
34									*	*	*				
32									*	*	*	*			
30									*	*	*	*			
28									*	*	*	*			
26									*	*	*	*			
24									*	*	*	*			
22									*	*	*	*			
20									*	*	*	*			
18								*	*	*	*	*	*		
16								*	*	*	*	*	*		
14								*	*	*	*	*	*		
12						*	*	*	*	*	*	*	*		
10						*	*	*	*	*	*	*	*		
8						*	*	*	*	*	*	*	*		
6						*	*	*	*	*	*	*	*		
4				*		*	*	*	*	*	*	*	*	*	
2				*	*	*	*	*	*	*	*	*	*	*	
Interval ...	18	21	24	27	30	33	36	39	42	45	48	51	54	57	60

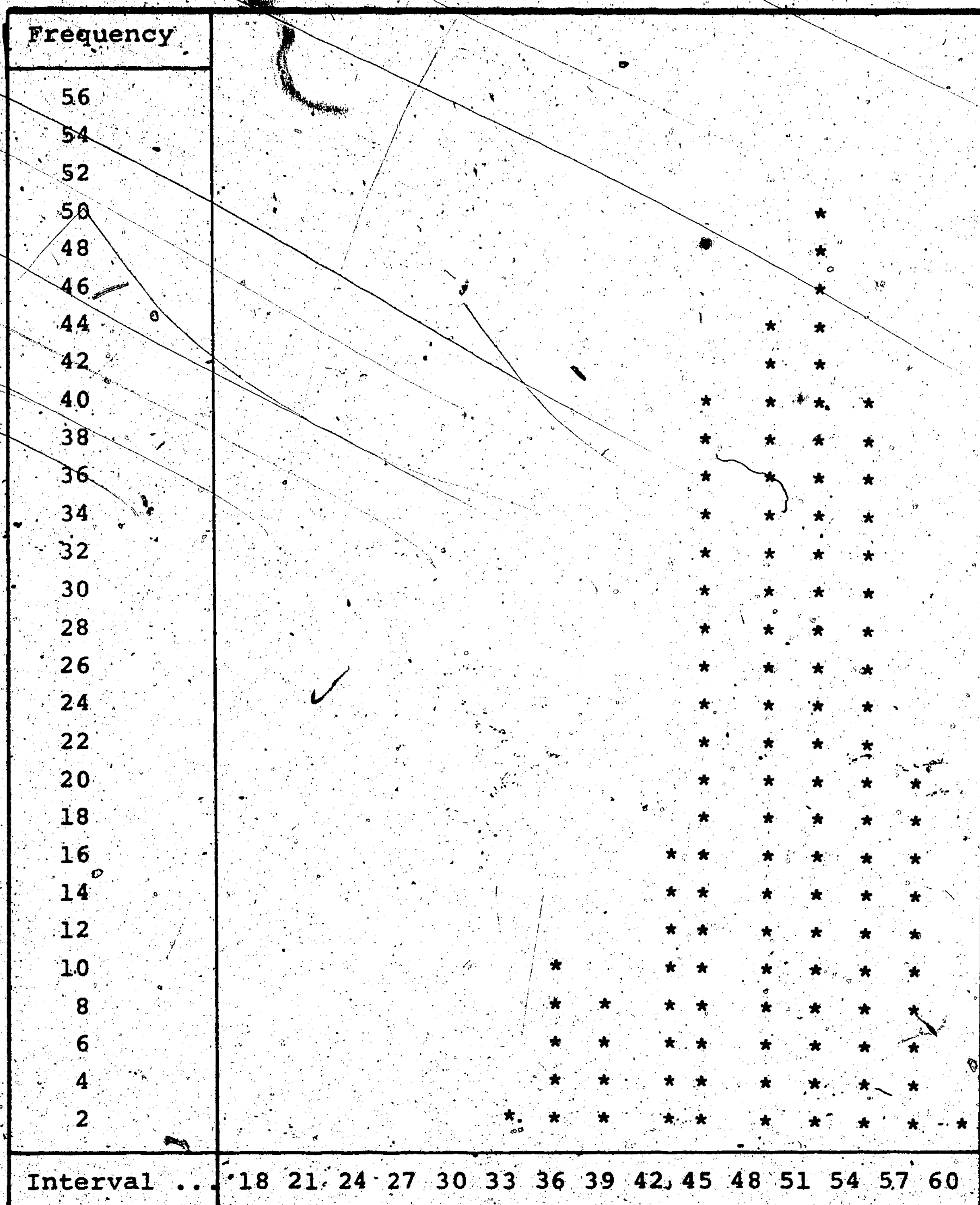
APPENDIX III

Grade Eight Frequency Distribution

Frequency															
56															
54															
52															
50												*			
48												*			
46											*	*			
44											*	*			
42										*	*	*			
40										*	*	*			
38										*	*	*	*		
36										*	*	*	*		
34										*	*	*	*		
32									*	*	*	*	*		
30									*	*	*	*	*		
28									*	*	*	*	*		
26									*	*	*	*	*		
24									*	*	*	*	*		
22									*	*	*	*	*		
20									*	*	*	*	*		
18									*	*	*	*	*		
16									*	*	*	*	*		
14									*	*	*	*	*	*	
12								*	*	*	*	*	*	*	
10								*	*	*	*	*	*	*	
8							*	*	*	*	*	*	*	*	
6						*	*	*	*	*	*	*	*	*	
4					*	*	*	*	*	*	*	*	*	*	
2		*			*	*	*	*	*	*	*	*	*	*	
Interval ...	18	21	24	27	30	33	36	39	42	45	48	51	54	57	60

APPENDIX IV

Grade Nine Frequency Distribution



APPENDIX V-A

Progressive Matrices
Point - Biserial Correlations
-Grades Seven Eight and Nine-

Test Item	Grade Seven	Grade Eight	Grade Nine
A 1	0.0	0.01	0.42
A 2	0.11	0.15	0.26
A 3	0.09	0.13	0.11
A 4	0.04	0.00	0.32
A 5	0.18	0.06	0.33
A 6	0.10	0.10	0.26
A 7	0.25	0.21	0.50
A 8	0.18	0.18	0.32
A 9	0.25	0.21	0.27
A10	0.21	0.32	0.38
A11	0.42	0.37	0.40
A12	0.50	0.29	0.30
B 1	0.02	0.14	0.27
B 2	-0.02	0.08	0.24
B 3	0.29	0.21	0.42
B 4	0.24	0.18	0.42
B 5	0.22	0.42	0.38
B 6	0.33	0.30	0.37
B 7	0.45	0.31	0.46
B 8	0.52	0.23	0.27
B 9	0.45	0.29	0.41
B10	0.50	0.31	0.40
B11	0.44	0.37	0.37
B12	0.51	0.42	0.43

APPENDIX V-B

Progressive Matrices
Point-Biserial Correlations
-Grades Seven Eight and Nine-

Test Item	Grade Seven	Grade Eight	Grade Nine
C 1	0.16	0.34	0.46
C 2	0.35	0.35	0.56
C 3	0.43	0.25	0.37
C 4	0.2[illegible]	0.22	0.32
C 5	[illegible]	0.08	0.54
C 6	0.38	0.44	0.39
C 7	0.44	0.48	0.47
C 8	0.47	0.41	0.43
C 9	0.35	0.29	0.28
C10	0.45	0.43	0.44
C11	0.36	0.36	0.31
C12	0.16	0.32	0.37
D 1	0.22	0.09	0.59
D 2	0.34	0.39	0.39
D 3	0.31	0.29	0.51
D 4	0.43	0.25	0.44
D 5	0.38	0.27	0.52
D 6	0.41	0.35	0.45
D 7	0.39	0.32	0.35
D 8	0.39	0.31	0.41
D 9	0.44	0.35	0.47
D10	0.49	0.37	0.50
D11	0.20	0.20	0.19

APPENDIX V-C

Progressive Matrices
Point-Biserial Correlations
-Grades Seven Eight and Nine-

Test Item	Grade Seven	Grade Eight	Grade Nine
E 1	0.42	0.29	0.46
E 2	0.43	0.44	0.50
E 3	0.52	0.44	0.37
E 4	0.58	0.54	0.50
E 5	0.53	0.57	0.56
E 6	0.49	0.47	0.47
E 7	0.42	0.40	0.40
E 8	0.37	0.49	0.47
E 9	0.24	0.40	0.41
E10	0.29	0.35	0.41
E11	0.09	0.08	0.21
E12	0.05	0.16	0.21

APPENDIX VI

Standard Error of Percentiles

(Age Norms)

Percentile / Age		25	50	75
12-0 - 12-5	S E	.8	.73	.8
	Interval	38.4-41.6	42.54-45.46	46.4-49.6
12-6 - 12-11	SE	.9	.8	.9
	Interval	39.2-42.8	43.4-46.6	47.2-50.8
13-0 - 13-5	SE	.77	.69	.77
	Interval	39.46-42.54	43.62-46.38	48.46-51.54
13-6 - 13-11	SE	.83	.75	.83
	Interval	41.34-44.66	45.5-48.5	49.34-52.66
14-0 - 14-5	SE	.83	.75	.83
	Interval	39.34-42.66	45.5-48.5	49.34-52.66
14-6 - 14-11	SE	.9	.8	.9
	Interval	42.2-45.8	45.4-48.6	49.2-52.8

Note: All intervals are at the 95% level of confidence

APPENDIX VII

Standard Error of Percentiles

(Grade Norms)

Percentile / Grade		25	50	75
7	SE	.63	.56	.63
	Interval	37.74-40.26	42.88-45.12	46.74-49.26
8	SE	.53	.48	.6
	Interval	40.94-43.06	45.04-46.96	49.04-50.96
9	SE	.6	.54	.6
	Interval	41.8-44.2	46.92-49.08	49.8-52.2

Note: All intervals are at the 95% level of confidence