University of Rajshahi **Rajshahi-6205** **Bangladesh.**

RUCL Institutional Repository **http://rulrepository.ru.ac.bd**

Department of Psychology PhD Thesis

1974

The Construction and Standardization of a Non-Language Test of Mental Ability for Secondary School Students (Class IX-X)

Shah, Amzad Hossain

University of Rajshahi

http://rulrepository.ru.ac.bd/handle/123456789/319

Copyright to the University of Rajshahi. All rights reserved. Downloaded from RUCL Institutional Repository.

The Construction and Standardization of a Non-Language Test of Mental Ability for Secondary School Students (Class IX-X)

THE CONSTRUCTION AND STANDARDIZATION OF A NON-LANGUAGE TEST OF MENTAL ABILITY FOR SECONDARY SCHOOL STUDENTS (CLASS IX-X)

A

dissertation

presented in fulfilment of the requirements

for the Degree of Doctor of Philosophy

in Psychology in the

University of

Rajshahi

Rajshahi University Library
Documentation Section
Document No. D
Date 9/3/88

by

Amzad Hossain Shah, M.Sc.

May, 1974.

Department of Psychology,
University of Rajshahi,
Bangladesh.

THE DISSERTATION IS APPROVED

BY

Professor Barrie Hopson, Ph.D.
Department of Psychology
University of Leeds
E_N_G_L_A_N_D

Professor Henry Dizney, Ph.D.
Department of Educational Psychology
University of Oregon
U_S_A

Professor Moslem Huda, Ph.D.
Department of Psychology
University of Rajshahi
B_A_N_G_L_A_D_E_S_H

ACKNOWLEDGEMENTS

I express my deep gratitude to Dr.Moslem Huda,Professor and Chairman, Department of Psychology, Rajshahi University, but for whose kind help and supervision the research work leading to this doctoral dissertation would not have been possible. I must thank Mr. A.H.Talukdar, Associate Professor in Statistics, Rajshahi University, for his valuable advice in the statistical analysis of the data. I thankfully acknowledge the help kindly extended to me by my colleagues of the Department of Psychology.

I feel grateful to the University authority for kindly providing me with the fund necessary for the smooth conduct of this research work. I express my gratitude to the Headmasters/ Headmistresses of the High Schools who have extended their sincere cooperation, and to boys and girls who have patiently taken the test. I thank Mr. Momtaj Ali of the Department of Geography for the cartographic assistance he has kindly given me.

The final form of the test was applied to another random sample (25% of the population) of students of class IX for standardization. The reliability (odd-even and test-retest) and the predictive validity were estimated. The factorial composition of the test was determined by the complete centroid method (Thurstonian). Norms were prepared for boys and girls in terms of a new scale developed on the assumption of normality of distribution of intelligence test scores.

Boys were found superior to girls in their performance on NLTMA. Science students did significantly better than Humanities students. Students of government schools did significantly better than those of non-government ones. The differential performances of these groups of subjects have been explained.

At the end an attempt was made to investigate whether the subjects would do better on NLTMA than on Raven's SPMT and Terman's NLMMT and whether the former correlated with the latter. These three tests were applied to another random sample of boys and girls of class X. Results showed that the performance was significantly better on NLTMA and that it correlated .74 and .64 with SPMT and NLMMT respectively. The significant difference between the mean performances may be ascribed to cultural factors. The significant correlation of NLTMA with SPMT and NLMMT may be regarded as an additional support for the validity of the former.

CONTENTS.

		Page
CHAPTER ONE	INTRODUCTION	1
CHAPTER TWO	REVIEW OF THE LITERATURE ...	7
CHAPTER THREE	CONSTRUCTION AND STANDARDIZATION...	25
CHAPTER FOUR	TEST ADMINISTRATION	33
	1. ADMINISTRATION I ...	33
	(i) Introduction ...	33
	(ii) Method	33
	(iii) Results	34
	2. ADMINISTRATION II ...	35
	(i) Introduction ...	35
	(ii) Method	35
	(iii) Results	36
	(iv) Discussion ...	46
	3. ADMINISTRATION III ...	48
	(i) Introduction ...	48
	(ii) Method	48
	(iii) Results	49
	(iv) Discussion ...	62
	(v) Norms	71
	(vi) Comparison of Mean Performances	74

		Page
CHAPTER FIVE	CORRELATIONAL STUDIES... ...	79
	1. ADMINISTRATION I ...	79
	(i) Introduction ...	79
	(ii) Method	79
	(iii) Results	80
	(iv) Discussion... ...	85
	2. ADMINISTRATION II ...	87
	(i) Introduction ...	87
	(ii) Method	87
	(iii) Results	88
	(iv) Discussion... ...	91
CHAPTER SIX	GENERAL DISCUSSION AND CONCLUSION	94
BIBLIOGRAPHY	... ...	101
APPENDIX		

CHAPTER ONE

INTRODUCTION

In this chapter the nature and purpose of the present study have been outlined and certain terms used throughout the report, explained.

The problem of the present study is to construct and standardize a non-verbal test of intelligence, named "A Non-Language Test of Mental Ability (NLTMA) for Secondary School students (Class IX-X)". A psychological test may be defined as an instrument designed for the quantitative assessment of psychological attributes of an individual. An attribute is some isolable characteristic of the individual, some dimension of structure or function along which the individual may be placed in some order (Nunnally, 1967,p.254). A psychological test is also an objective measure of a sample of behaviour (Anastasi, 1963, p.21). In other words, a psychological test is an objective and standardized measure intended to evaluate a certain aspect or attribute of the human ability by means of a sample of performance

or behaviour. A psychological test is, then, essentially a combination of a number of items referring to the activities involving the psychological attributes of the individual for whom the test is designed. Each of these items acts as the stimulus to which the individuals respond in some specified (sometimes, unspecified) ways. That is why a psychological test is defined as a pattern of stimuli selected and organized to elicit responses which will reveal certain psychological characteristics in the person who makes them (Mursell,1947, p. 1). For the purpose of the present study we have followed the definition of a psychological test given by Professor L.L.Thurstone, who defines it " as the task, together with the method of appraising it, which defines an ability" (Thurstone, 1947, p. 62).

The objectivity of a psychological test implies that the administering and scoring procedures be framed up in

such a manner that the influence of the subjective judgement is reduced to a minimum. In other words, the objectivity of a psychological test denotes the degree of its freedom from errors on account of personal feelings and bias, its ability to reflect the facts of the problematic situation irrespective of the personal equation of the agent. A psychological test is said to be standardized when a uniformity of procedure in administering and scoring the test is attained (Anastasi, 1963, p. 23). The entire process of try-outs, statistical analysis and evaluation on the basis of accepted criteria is called test standardization (Freeman, 1960, p. 10).

After having defined a psychological test, we may now delimit the concept of intelligence, which the present test is intended to assess. Factorial studies by prominent psychologists like J.P.Guilford and L.L.Thurstone have

revealed that the spatial and perceptual abilities form an essential feature of the general intellectual ability (Anastasi, 1963, p.345 ; Freeman, 1960, p. 78 ; Guilford, 1967, pp. 186-187 and 433-434). The present test has been designed to evaluate mainly the spatial and perceptual abilities. The spatial ability requires a subject to manipulate objects imaginally in space and the perceptual ability involves the seeing of spatial and perceptual relations with rapidity and accuracy.

A non-language test of mental ability is one in which an individual requires the use of paper and pencil but no knowledge of reading and writing to respond to the test items. Some tests of this kind have been developed and standardized in countries abroad. But no statistical evidence in support of the culture-freeness of these tests is available. These tests are composed of items many of which

are unfamiliar to us. Again, most of the relevant tests provide no information regarding the coefficient alpha (internal or interitem consistency reliability), the importance of which for a test of mental ability has been discussed by prominent psychologists (Guilford, 1954, pp. 380-383 ; Nunnally, 1967, pp. 196-198). If the test manual, as Nunnally puts it, either gives no reliability information or states a reliability coefficient without saying how it was determined, the test should be viewed with suspicion (Nunnally, 1959, p.112). As Guilford has commented -" The internal consistency reliability is one of the major classes of reliability, operationally defined."(Guilford, 1954, p.374). The interitem consistency reliability coefficient provides a measure of both equivalence (as in split-half methods) and homogeneity (Anastasi, 1963,p.122). Anastasi further remarks that the test scores are less ambiguous when derived from relatively homogeneous tests (Anastasi, 1963, p. 109). Freeman comments that the higher

Appendix to Chapter One.

Under the educational system of Bangladesh a student at the level of Class IX only can opt for a specific course of study keeping in view the future profession he intends to take up. But unfortunately for our students no tests are available to help them choose a particular course of study suited to their ability.

Following the emergence of Bangladesh the Government and the educationists of the country have realized the need for constructing various tests so that students can be sorted out for various courses of studies according to their interests and aptitudes.

The present test is the first of its kind in this country, and there is at the moment no work with which it can be compared. However, the present work has been undertaken to help students test their ability to do well in such fields as drafting, dress designing, architecture, art, die-making and decorating, and also in the Air Force.

The test may be used in Bangladesh for purposes of identifying at an early level of education those students who have aptitude for studies referred to above but who do not fare well on verbal tests. The test may also be used in combination with verbal tests for having a global estimate of the general mental ability of our students.

the test's reliability coefficient, the smaller will be the error of measurement and therefore the greater the predictive value of the test (Freeman, 1960, p. 18). It may likewise be noted that the higher the coefficient alpha of a test, the higher will be its homogeneity and therefore the less the ambiguity of the test scores.

Keeping in view the aforesaid disadvantages of most of the available tests, the present one has been designed to suit our needs with a high coefficient alpha so that it may yield unambiguous scores. It is a group test of mental ability intended to evaluate mainly the intelligence of individuals of our country.

Inasmuch as it consists of items divested of language factors, it may also be used to assess the intelligence of those who are physically handicapped or illiterate, or are poor in their performance on verbal tests.

CHAPTER TWO

REVIEW OF THE LITERATURE

In this chapter an attempt has been made to survey the literature on the construction and standardization of some tests of mental ability relevant to the present study. This has been done with the purpose of gaining insight into the fundamental characteristics of these tests and the disadvantages associated with most of them. The construction and standardization of a test of mental ability essentially involves, inter alia, the fulfilment of the following fundamental characteristics : (1) The nature of items used (content validity), (2) reliability (Test-retest, split-half and parallel-form), (3) coefficient alpha (interitem consistency reliability),(4) validity (concurrent ,predictive and factorial), and (5) norm. In the following pages the discussion will, therefore, be limited mainly to the aforesaid points. Besides, experimental evidences will be presented at the end to examine the culture-fairness of the tests.

Non-Language Group Tests of Intelligence.

The nature of items used.

The Army Beta Examination was the first non-language group test of intelligence constructed for testing foreign-speaking and illeterate soldiers in the army during world war I. The original Army Beta Examination consisted of 5 subtests: (1) Familiar mazes ; (2) cube analysis and counting; (3) symbol-digit combinations; (4) pictorial completion, and (5) geometrical construction. The revised Beta Examination comprises 6 subtests : (1) Mazes ; (2) symbol-digit substitution ; (3) pictorial completion ; (4) paper form board ; (5) pictorial absurdities, and (6) perceptual speed.

The Australian Council for Educational Research Non-verbal Test is designed for the measurement of intelligence of children from 10 to 14 years. The test contains the following subtests : (1) Similarities test ; (2) geometrical relations ; (3) analogies test ; and (4) time sequence. The number of items in the test is 56.

The Non-Language Multi-Mental Test is designed for the assessment of intelligence of subjects of a grade range from 3 to 8. The test is available in two parallel forms - Form A and Form B. Each of the two forms comprises 60 items. Every item employs the drawings of 5 objects,

the four of which may be grouped together on the basis of some common relations but the fifth one is different. The subject has to identify the one which is different from other objects. The items of the test are designed to measure " the ability of an individual to recognize and utilize relationships not among verbal symbols, but rather among pictorial symbols." (Freeman, 1960,p.252; Terman et al., 1942, p.1).

The Pintner Non-Language Scale was originally developed to measure the intelligence of deaf children. In the revised form, the scale consists of two forms- Form K and Form L, designed for the assessment of intelligence of children from grades 4 to 9. The scale contains 6 subtests : (1) Figure dividing ; (2) reverse drawing ; (3) pattern synthesis ; (4) movement sequence ; (5) manikin, and (6) paper folding.

The Pattern Perception Test employs only a single type of material, designed for the assessment of intelligence of adults. The test contains materials which are complex in nature and require subtle perception and reasoning for their solution. This test comprises 60 items, each of which contains a row of 5 designs. The subject has to identify the four designs which form a pattern and to cross out the one which seems to be inappropriate. The eight sets of items which constitute the test are arranged in order of increasing difficulty.

Raven's Standard Progressive Matrices, divided into 5 sets (sets A,B,C,D and E) of 12 items each, is designed to assess the intellectual ability of children as well as adults. The matrices test comprises 60 items, arranged in order of increasing difficulty. The items are designed to test the ability of an individual to comprehend meaningless figures presented for his observation, see the relations between them, conceive the nature of the figure completing each system of relations, and thus develop a systematic method of reasoning.'

Cattell Culture-Fair Intelligence Test (Scale 3) is a non-language test constructed to provide a measure of general intellectual ability of High School pupils and adults. The test comprises two parallel forms - Form A and Form B. Each of the forms consists of 50 items, divided into 4 subtests such as : (1) Series, involving the ability of an individual to identify the item that would complete the series; (2) classification, tapping the ability of a subject to mark the one item, in a row of 5 figures, that does not belong with the others; (3) matrices, requiring an individual to identify the item which would correctly complete the given matrix, and (4) conditions, measuring the ability of an individual to insert a dot in one of 5 alternative designs so as to meet the same conditions indicated in the sample design.

Rajshahi University Library
Documentation Section
Document No. D-1044
Date. 2.12.88

The Figure Reasoning Test is a non-verbal test of intelligence constructed by J.C.Daniels on the principles of Raven's Progressive Matrices. The test contains 45 items, each of which has a 3x3 square of a set of diagrams with the last one missing. The testee has to identify the missing one from among 6 choices. The order of items is a compromise between the logical sequence and the degree of difficulty.

The Chicago Non-verbal Examination is a non-language test designed to assess the intelligence of subjects from 6 years to adulthood. This test consists of 7 subtests : (1) Block counting; (2) paper form board ; (3) matching figures ; (4) picture sequence ; (5) picture completion; (6) symbol-digit substitution, and (7) picture absurdities.

Non-verbal Tests I and II were designed by J.W.Jenkins and D.M.Lee for the assessment of intelligence of subjects from 10 to 13 years. The test I is composed of 5 subtests. Jenkins uses the mathematico-logical principles of classification, ordered sequence and analogy and applies them to non-verbal pattern materials. The subtests 1 and 3 use the device of like and unlike as a test of the principle of classification; the subtests 2 and 4 utilize the familiar device of ordered sequence and the subtest 5 uses analogy. The test II employs the same logical principles as the test I in regard to the construction of items.

The Goodenough Draw-a-Man Test is a non-language test designed to evaluate a child's intellectual ability by means of the drawing of a man. This test is intended for use with subjects from 3.5 to 13.5 years. The test has been extensively revised, and the revised form goes by the title of Harris-Goodenough Test of Mental Maturity. In the original test, the subject is simply asked to draw the picture of a man as best as he can. He is further instructed to work carefully and take his time. In the original test as well as in the revised one, an emphasis is laid on the child's accuracy of observation and on the development of conceptual thinking. 73 items were selected on the basis of age differentiation, relation to the total score and relation to the group intelligence test score. In the revised form, the subject is asked to draw the picture of a woman and of themselves.

Davis-Eells Test of General Intelligence is a test of general intelligence or problem-solving ability. The test is designed to present situations and problems in forms which are covered by experiences common to all children in the group for whom they are intended. The test is composed of two forms : (1) Primary for grades I and 2, and (2) elementary for grades 3 to 6. There are 88 items in the test. The best-way items are 29; the probabilities items, 29; the picture analysis items, 22, and the money items, 8.

Reliability.

The Revised Army Beta Examination has an odd-even reliability coefficient of .98 and a test-retest reliability coefficient of .77. The A.C.E.R. Non-verbal Test provides a split-half reliability coefficient of .85 but no information about the test-retest.

The Non-Language Multi-Mental Test provides for grades 3 to 8 reliability coefficients of .86 for Form A, .90 for Form B and .94 for both forms combined. Reliability coefficients for each of the several grades taken separately vary from .66 to .74 for each form taken alone. The estimated reliability for both forms combined is .80 (Freeman, 1960, p. 253).

The Pintner Non-Language Scale has a split-half reliability coefficient of .85 for Form K and .89 for Form L. The test provides no information regarding the test-retest reliability coefficient. The Pattern Perception Test has a test-retest reliability coefficient of .80 to .90, but it provides no information regarding the split-half reliability coefficient.

The Standard Progressive Matrices provides a test-retest reliabiliy coefficient varying with age from .83 to .93. Retest reliability in groups of older children and adults, moderately homogeneous in age, has been found to vary approximately between .70 and .90. But the test has no reference to the split-half reliability coefficient.

Cattell Culture-Fair Intelligence Test has a test-retest reliability coefficient varying between .84 to .94. Rodd found the reliability to fall as low as .53. The split-half reliability coefficients for three undergraduate groups (100, 155 and 212) were estimated to be .82, .91 and .95 respectively. The split-half reliability coefficient for the two forms (A & B) combined varies from .70 to .92.

The Figure Reasoning Test has a split-half reliability coefficient of .96 and a test-retest reliability of .70 after an interval of one year. The Chicago Non-verbal Examination provides a reliability coefficient varying from .80 to .90. The test does not have any reference to the method of determining reliability.

Jankins' Non-verbal Test 1 has a split-half reliability coefficient of .95. The reliability coefficient of Test II is .92 and the two parallel forms correlate .87. But the test does not provide any information regarding the test-retest reliability coefficient.

Goodenough Draw-a-Man Test has a test-retest reliability coefficient of .68 after an interval of 7 days and a split-half reliability coefficient of .89. Davis-Eells Test of General Intelligence provides split-half reliability coefficients varying from .68 (grade I) to .81 (grade VI). The test-retest reliability coefficients range from .72 (grade II) to .90 (grade IV).

As noted above, some tests do not provide any data on reliability (test-retest) statistics. Hence one cannot be sure if the psychological functions assessed by the tests maintain stability with passage of time. Again, some tests provide no information regarding the split-half reliability statistics. It casts doubt as to whether or not the tests provide a measure of equivalence, or of adequate sampling of items. Besides, some tests have no reference as to the method of determining the reliability statistics.

While developing the present test, an attempt has been made to eliminate as far as possible the disadvantages referred to above.

Coefficient alpha.

Most of the tests already mentioned provide no information concerning the coefficient alpha (interitem consistency reliability coefficient) which measures both equivalence and homogeneity (vide Chapter 1 at p.5). To ensure homogeneity of test items and remove ambiguity of test scores, there is greater need for reliance on the coefficient alpha.

Hence, inthe present test, the coefficient alpha has been rigorously determined and reported.

Validity.

The validity of the Army Beta Examination is provided in terms of the correlation between its point scores and the S-B (1916) MA's, the correlation being .78 ± .02. The validity coefficient determined by the correlation between its point scores and Otis Self-Administering Test of Mental Ability scores amounts to .77 ± .02. Thus, the test provides information regarding the concurrent validity. But data on the predictive validity and the factorial validity of the test are not provided.

The validity coefficients of the A.C.E.R. Non-verbal Test are provided in terms of the correlations between its scores and (1) S-B scale, (2) Otis Self-Administering Test of Mental Ability, and (3) Tasmanian Educational Department's General Ability Test. These are respectively .63, .74 and .68. The test thus provides statistics on the concurrent validity, but no information regarding the predictive validity and factorial validity is available.

The Non-Language Multi-Mental Test provides the validity coefficients in terms of its correlations with CA and MA, the correlations being respectively .65 and .67. No data on the predictive validity and the factorial validity of the test are provided.

As the criteria for validity of the Pintner Non-Language Scale, Pintner used the following : (1) The increase in mean scores in successive age groups ; (2) correlation with his verbal test, and (3) similarity between his test and those devised by factor analysis. The correlations with his verbal test for homogeneous age groups indicating the concurrent validity are found to be in the .60's. But no data on the predictive validity of the test are available.

The Pintner scale is designed "to provide relatively independent measures of the spatial factor, perceptual ability (visual) and reasoning (without use of language)" (Freeman, 1960,p.250). This scale utilizes some of the factors presumably isolated by the adherents of the group-factor theory of intelligence. But Pintner does not attempt to indicate which of the subtests measure the presumed factors. Hence, Pintner's assertion that his scale taps the aforesaid factors need to be corroborated by factorial studies.

The validity coefficients of the Pattern Perception Test were determined by its correlations with job ratings in the British Navy and Army. These correlations ranged from .30 to .47. Its correlations with other tests varied from .43 for 67 medical students to .73 for a random sample of 597 men in the British Army. The test provides no information regarding the factorial validity.

Raven's Standard Progressive Matrices correlates with both verbal and performance tests of intelligence. These correlations indicating the concurrent validity of the test vary between .40 and .75, and have a tendency to be higher with performance than with verbal tests. The predictive validity of the test has been determined by its correlation with the academic criteria and found to run somewhat lower than those of the usual verbal intelligence tests. The test has a correlation of .86 with Terman-Merrill Scale.

Factorial studies suggest that the matrices tests largely measure a general factor (a g-saturation of .82) as well as a small spatial factor. Raven interprets this general factor as being essentially the same as Spearman's eduction of relations and correlates. But "spatial aptitude, inductive reasoning, perceptual accuracy, and other group factors also influence performance" on this test (Anastasi,1963,p.262).

Cattell Culture-Fair Intelligence Test correlates .85 and .84 with the Stanford-Binet Scale and Weschler-Bellevue Scale respectively. The C.F.test correlates .55 with the Army Alpha Examination Modified Form 9. These intertest correlations indicate the concurrent validity of the test. But "data on the concurrent and predictive validity in terms of non-test criteria are virtually non-existent" (Anastasi,1963, p.261). The "g" correlations of the culture-free subtests have been found to vary between .78 and .83.Bajard has found the g-saturations to range in different samples from .71 to .84 (Cattell,1959,p.7).

Daniels' Figure Reasoning Test provides no information regarding the statistics on the concurrent and predictive validity. Items of the test have been selected on the basis of their correlations with the criterion with rejection of those correlating less than .25. But the criterion used is not defined. The factorial analysis suggests a general factor saturation of .85 with an additional non-verbal factor.

The criteria for validity of the Chicago Non-verbal Examination are the following : (1) Correlation with CA ; (2) known groups (in this case, mentally retarded) ; (3) symmetrical bell-shaped distribution, and (4) correlations with other tests. Correlations between the test scores and CA vary from .57 to .81. The authors provide a satisfactory distribution of scores, which closely approximates the symmetrical curve. No data on the factorial validity of the test are provided.

Jenkins' Non-verbal tests correlate .81 with Raven's Progressive Matrices. This correlation represents the concurrent validity of the test. The authors provide no information concerning the predictive validity and factorial validity of their tests.

The earlier form of Goodenough Draw-a-Man Test correlates with other tests such as the S-B scale. These correlations indicating the concurrent validity of the test range from .41 to .80. The test provides no information regarding the predictive validity.

Correlations between the Draw-a-Man IQ and scores on other tests of known factorial composition indicate that, within the ages covered, the Draw-a-Man test correlates the highest with tests of reasoning, spatial aptitude and perceptual accuracy. But factorial studies need to be carried out to corroborate the above indication.

Davis-Eells Test of General Intelligence has been found to correlate with Otis Quick-Scoring Test of Mental Ability and standardized school achievement tests. These correlations have been found to vary respectively between .39 and .66, and .24 and .42. These correlations may indicate the concurrent and predictive validity respectively.

The authors utilize psychological interviews and psychological analysis rather than factor analysis to determine the psychological functions sampled by the situations and problems presented in the test. Hence, factorial studies need to be conducted to make the factorial composition of the test stand revealed.

As referred to above, some of the tests provide no information regarding the predictive validity. The predictive validity of a test indicates its effectiveness in predicting future performance. Again, some tests do not provide any data on the factorial validity. The factorial validity indicates the extent to which a test measures a theoretical construct or trait.

Hence the predictive validity and factorial validity of the present test have been rigorously estimated and reported.

Norms.

Norms of the Army Beta Examination are available in terms of the S-B (1916) mental age equivalents , the range being from MA 6 years and 3 months to 16 years and 8 months. Norms of the A.C.E.R. Non-verbal Test are provided in terms of IQ's and those of Non-Language Multi-Mental Test, in terms of mental age equivalents and grade equivalents.Norms of the Pintner Non-Language Scale are available in terms of mental ages and percentiles. Total scores on this test are expressed as deviation IQ's with an SD of 16.

The Pattern Perception Test does not provide any data on the norms. This may indicate that this test is still in the process of standardization. The Standard Progressive Matrices provides norms in terms of percentile points calculated from the natural sources of 735 Colchester children. Norms on this test are also provided in terms of percentile points calculated from the natural sources of 3,665 militiamen and 2,192 civilians.

Norms for the Cattell Culture-Fair Test are available in terms of percentiles and IQ's, and those for the Figure Reasoning Test are provided for ages 10 - 16 and adults in terms of z-scores. But the Chicago Non-verbal Examination provides no data on the norms. This may mean that the test is still in the process of standardization.

Jenkins and Lee provide norms for the Non-verbal Tests I and II over the age range from 10-0 to 12-11 in terms of normalized scores with a mean of 100 and an SD of 15. These norms are based on a mixed sample of boys and girls. Norms for the Goodenough Draw-a-Man Test are available in terms of deviation IQ's with a mean of 100 and an SD of 15, but no normative data for the Davis-Eells Test of General Intelligence are provided. This may cast doubt as to whether or not the test is in the process of standardization.

From the foregoing review of the literature on the construction and standardization of some relevant tests of intelligence, the following points stand revealed : (1) In some tests, data on validity and reliability statistics are wanting or too meagre to be of any significant use for purposes of testing . (2) In some tests, data on the reliability statistics are available but no information regarding the factorial validity of the tests has been provided.(3) Most of the tests provide no information concerning the coefficient alpha(interitem consistency reliability). In other words, no evidence has been put forward to show that they consist of homogeneous items. Hence, we are not sure if the test scores obtained from their administration are free from ambiguity. (4) Norms have nearly all been provided in terms of normalized z-scores, percentiles and MA equivalents. In some tests, normative data are not available.

The disadvantages of percentiles and other scales in use have been referred to in Chapter IV.

As regards complete culture-fairness of the existing non-verbal (non-language) tests of intelligence, evidences are not conclusive (Guilford, 1967, pp.408 - 409). The study of P.E. Vernon with non-verbal tests like Progressive Matrices, the non-verbal items of Terman-Merrill, and so on may be cited in corroboration of the above fact (Butcher, 1968, pp. 256 - 257). Cattell himself failed to show the culture-fairness of a non-verbal test, supposedly culture-free, in connection with his studies on Leicester 10-year-old children (Hunt, 1961, p.338). As Anastasi has commented — "The culture-free tests have been administered in several European countries, in America and in certain African and Asiatic countries. Norms tended to remain unchanged in cultures moderately similar to that in which the tests were developed ; in other cultures, however, performance fell considerably below the original norms" (Anastasi, 1963, p.261). Tsakalos found Moray House arithmetic test No.36 to be culturally fairer than Jenkins' non-verbal intelligence test No.1 (Butcher, 1968, p.269). Ortar in Israel failed to obtain evidence in support of the culture-fairness of non-verbal tests in her studies on high-status Israeli children and oriental immigrants. Davis, Lesser and French also could not show their own non-verbal tests to be completely culture-fair. Similar

findings were also arrived at by Lesser, Fifer and Clark in connection with their study regarding the culture-fairness of non-verbal tests (Butcher, 1968, pp. 253 - 258). Anastasi comments on Davis-Eells Test of General Intelligence that this test appears to have sacrificed predictive validity without eliminating cultural bias (Anastasi, 1963, p. 268). The presence of the cultural influence in the Goodenough Draw-a-Man Test, supposedly culture-free, was demonstrated by Havighurst et al. in their study on the North American Indians (Freeman, 1960, p.268).

In view of the above findings, it may be concluded that the cultural fairness has yet to be achieved in respect of the construction of non-language tests. Considering what have been said above, this study was designed to construct and standardize a non-language test of mental ability which may be useful for this country and free from the aforesaid disadvantages.

CHAPTER THREE

CONSTRUCTION AND STANDARDIZATION

In this chapter an attempt has been made to present an outline of the method followed in constructing and standardizing the present test (NLTMA).

Procedure.

A universe of 230 items was first composed, and these were classified into four subtests : Test A (55 items), Test B (55 items), Test C (70 items), and Test D (55 items). These were respectively called Figure matching test, Sign patterning test, Series completion test , and Pattern synthesis test according to the mode of working out the correct solutions for each of the test-samples.

The test consisting of these four subtests was then administered on a sample of 20 boys and girls reading in classes from VI to X. In the light of the data of this preliminary try-out, some items were modified without reducing their number.

This modified form of the test consisting of 230 items was then applied to a different random sample (50% of the population) of class IX pupils of 11 High Schools (both government and non-government) of Rajshahi Town in Bangladesh. The sample was drawn by means of a table of random numbers (Dixon et al., 1951, pp.290-294) and was statistically tested (normality test and χ^2-test) for randomness and representativeness. Item analysis was then performed by the Flanagan-Kelley method

(Flanagan, 1939, pp.674-680; Garrett, 1960, pp.365-367; Guilford, 1954, p.428 ; Kelley, 1939, pp. 17-24). The proportion (p) of right answer in the top 27% and in the bottom 27% of cases was calculated, leaving aside the middle 46% of cases. The difficulty index of each item was given by the proportion (p) (or, % of passes) of the individuals providing the right answer. The proportion of correct responses was then adjusted for chance success by using the formula,

$$p_c = \frac{R - \frac{W}{k-1}}{N - U} \qquad \ldots \qquad (1)$$

(Garrett, 1960, p. 364 ; Guilford, 1954, p. 420). The difficulty index (p_c) of each item was then obtained by averaging its p_c's in the top and bottom groups.

Item variances were then determined. The variance of an item is given by the product of its p-value (proportion of right answers) and q-value (proportion of wrong answers) (Garrett, 1960, p.363 ; Ghiselli, 1964, p. 65 ; Weatherburn, 1962, p. 116). As the maximum variance is given by the p-value of .5, items having p-values close to .4 to .6 (optimum difficulty) were selected for the final form of the test, though items having extreme p-values in either direction were also retained by way of basal and ceiling ones.

Item validity (or the validity index of the item) was then determined in terms of biserial correlation coefficient. The Flanagan method considered to be the most efficient was used for estimating the

biserial correlation coefficient. The choice of 27% in each tail is justified by Kelley's demonstration that, with this tail proportion, the coefficient of biserial correlation has been found to be the most sensitive (Guilford, 1954, pp. 428-429 ; Kelley, 1939, pp.17-24). The discriminative power of an item is, as Garrett has commented, most accurately determined when item analysis is based on the top and bottom 27% rather than some other percentage of the distribution (Garrett, 1960, p. 367). Item validity indicates the discriminative power of an item and the biserial correlation coefficient determines the extent to which an item bears correlation with the total score. The biserial correlation was read directly from a table of normalised biserial correlation coefficients (Garrett, 1960, p. 365 ; Tate, 1955, p. 364). The biserial correlation was used instead of point biserial correlation on the assumption of a continuous normal distribution of trait down from dismal failure to passing with the greatest ease (McNemar, 1949, p. 173). The design of item analysis has been shown in Tables I and II.

TABLE I

Item Analysis (Frequency Table)

#	Right R	Wrong W	Unattempted U	N - U

TABLE II

Item Analysis .

#	Difficulty Index (corrected for chance success) p_c	Item variance pq	Validity index r_b (Biserial Correlation Coefficient)	Remarks

While selecting items for inclusion in the final test, three points were kept in view : (1) the difficulty index of an item, (2) item variance and (3) item validity. Emphasis was laid more on the item validity than on the other points. As a rule, an item validity of .20 or more is regarded as satisfactory (Garrett, 1960, p. 368). The length of the test was determined by the size of coefficient alpha and, hence, of the inter-item consistency reliability, the objective being to have a homogeneous test.

The coefficient alpha of each test-sample and also of the whole test was calculated by using the K-R formula 20,

$$r_{kk} = \frac{k}{k-1}\left(1 - \frac{\Sigma pq}{\sigma_y^2}\right) \quad \ldots \quad \ldots \quad (2)$$

(Ghiselli, 1964, p. 286 ; Guilford, 1954, p. 380 ; Nunnally, 1967, p. 196; Richardson et al., 1939, pp. 681-687). The coefficient alpha for the NLTMA was desired to be ≥ 0.95. The number of items was increased until the desired level of coefficient alpha was obtained. That the reliability of a test increases with increase in the number of items can be algebraically shown by using the formula,

$$r_{kk} = \frac{k\,\bar{r}_{ij}}{1+(k-1)\,\bar{r}_{ij}} \quad \ldots \quad \ldots \quad (3)$$

(Nunnally, 1967, p. 193).

From equation (3), we have , $r_{kk} = \dfrac{1+(k-1)\,\bar{r}_{ij} - 1+\bar{r}_{ij}}{1+(k-1)\,\bar{r}_{ij}}$

$$\text{or,} \quad r_{kk} = 1 - \frac{1-\bar{r}_{ij}}{1+(k-1)\,\bar{r}_{ij}} \quad \ldots \quad (4)$$

From equation (4), it is evident that the right hand side increases as k, the number of items, increases. It is also clear that the reliability can never be greater than 1.00, however infinitely large the number of items may be.

Thus, determining the size of the coefficient alpha and the appropriate test-length, items were arranged in the final form in order of increasing difficulty and not in order of item validity.

The final form of the NLTMA was then administered to another random sample consisting of 25% of class IX pupils of 13 High Schools (both Government and non-government) of Rajshahi Town in Bangladesh for purposes of standardization. The sample was drawn by means of a table of random numbers (Dixon et al., 1951, pp. 290-294). The randomness and representativeness of the sample were statistically tested (normality test and χ^2-test).

The NLTMA was then subjected to relevant statistical tests to see if it would satisfy the criteria for a power test. For a power test, $M_w = M_x$, $\sigma_w = \sigma_x$, $\sigma_u = 0$ (since $u=0$), and $\frac{\sigma_u}{\sigma_x}$ is either very small or equal to zero (Gulliksen, 1950, pp. 231-233).

Since the NLTMA was designed to be a power test, the split-half (odd-even) technique (Guilford, 1965, p. 451 ; Guilford, 1954, p.376) was used to determine the reliability of the test by means of the formula,

$$r_{oe} = \frac{r_{ot}\,\sigma_t - \sigma_o}{\sqrt{\sigma_t^2 + \sigma_o^2 - 2r_{ot}+\sigma_o\,\sigma_t}} \quad \ldots \quad \ldots \quad (5)$$

(Guilford, 1954, p. 377 ; Mosier, 1941, pp. 407-408). The reliability of the whole test was then estimated by using the Spearman-Brown Prophecy formula, which reads

$$r_{tt} = \frac{2r_{oe}}{1+(2-1)r_{oe}} = \frac{2r_{oe}}{1+ r_{oe}} \quad \ldots \quad (6)$$

in this case. The reliability of the NLTMA was also calculated by the test-retest method on the lapse of a period of two months following the third administration.

The predictive validity of the NLTMA was determined by calculating the correlation between the test scores and the marks obtained in the school annual examination which was held a few days after the third administration of the test. In order to determine the factorial validity of the test, the complete centroid method of Prof.L.L.Thurstone was employed. For this purpose, the PM correlations between the test-samples, as calculated by the method of scattergram in virtue of the formula,

$$r = \frac{\frac{\Sigma x' y'}{N} - c_x c_y}{\sigma'_x \sigma'_y} \quad \ldots \quad \ldots \quad (7)$$

(Garrett, 1960, p. 138), were used in the correlation matrix. Factor analysis was, however, carried out on the test scores obtained by the standardization sample which was uniform in respect of age and academic level but not of sex on the ground that in multivariate analysis too much experimental control is not necessary (Guilford, 1968,p.218). The orthogonal rotation was then performed on the centroid factors in order to render them psychologically meaningful. The significance of the factor loadings on the subtests was statistically tested by calculating the standard error of a loading r by virtue of the formula,

$$SE_r = \frac{(1-r^2)\sqrt{n}}{\sqrt{N\,(n-s+1)}} \quad \ldots \quad (8)$$

(Vernon, 1950, p. 130). An interpretation of factors was then presented.

To prepare the norm, a new scale (called R-scale) was developed.

The mean performances of boys and girls, of students of Science and Humanities group, and of students of government and non-government High Schools on the NLTMA were compared, the main purpose being to see whether they differ significantly one from another.

Appendix to Chapter Three.

The present test is designed to assess mainly the spatial and perceptual abilities of an individual. In order that the test may serve this purpose, the four subtests similar to those used by several psychologists in their non-language tests of mental ability were selected (Anastasi,1963,pp.250-268; Cattell et al.,1959, pp.5-8; Cronbach,1960,pp.276-281; Freeman,1960,pp.245-259; Guilford,1967,pp.91-94, 100-102,143-145; Guilford,1968,pp.229-230;Terman et al.,1942,p.1). The subtests used by eminent psychologists like Bennett, Cattell, Penrose, Pintner, Raven,Terman and Thurstone proved to be valid measures of spatial and perceptual abilities. But since the items of these subtests were considered to be foreign to our culture and hence unfamiliar to our subjects, these were modified to bring them within the range of familiarity of our subjects irrespective of their religious and social differences.

While preparing the items constituting the subtests, the basic principles underlying the subtests used by those psychologists were followed. To test if our subjects are more familiar with these newly drawn items, those having validity indices of .20 or more were administered on a sample on which Raven's SPMT and Terman's NLMMT were also applied one after another (vide pp.79-93 of this thesis). Results showed that our subjects did significantly better on this test than on the latter (vide pp.85 and 91 of this thesis). This indicates that our test-samples better suit our culture and that our subjects are more familiar with them. The significant correlations between our test and Raven's SPMT and Terman's NLMMT (vide pp.85-86 and 92 of this thesis) may also be taken to mean that the test is a valid measure of spatial and perceptual abilities (vide pp.62-64 of this thesis). Results of item analysis show that the test is homogeneous (vide p.48 of this thesis).

CHAPTER FOUR

TEST ADMINISTRATION

In this chapter administrations of the present test have been reported and the results obtained therefrom, presented. Besides, some discussions about the main features of item analysis and factor analysis have been presented.

ADMINISTRATION I

INTRODUCTION

The administration I was undertaken to detect the defects, if any, in the construction of items of the test.

METHOD

Subjects.

An unselected sample of 20 students reading in classes VI to X of 4 High Schools (government and non-government) of Rajshahi Town in Bangladesh was used as subjects in this preliminary try-out. This sample consisted of 70% boys and 30% girls, of which 60% belonged to Humanities group and 40%, to Science group.

Tests used.

Non-Language Test of Mental Ability consisting of 230 items.

Procedure.

The test was administered on a group and not individually. Desks were placed widely apart in a suitable room. Subjects were requested to sit one behind another. Care was taken to see that they felt at home and

enjoyed the testing situation. Each of them was provided with a pencil and a rubber eraser. They were requested to fill up the blanks at the top of the answer sheet. Then the test booklet was distributed to each of them. Instructions (vide Appendix) for taking the test were read out in Bengali (English versions of instructions are also provided), and, in order to make these clear to the subjects, practice items were drawn on the blackboard and the working out of their solutions, explained. Praise and encouragement were given during administration to get them motivated. Subjects were asked to take all the time they required and not to be in a hurry. They were also asked not to go to the next page until they had completed the previous one. Subjects were further instructed to verbalise the defects, if they found any, in the drawing of items.

There was one scoring key for each of the four subtests. The key was placed over the answer sheet in such a way that the markings could be seen through the holes. The keys were so made as to block responses other than the right ones. The summation of marks obtained in the subtests constituted the total score on the test.

RESULTS

On the basis of the results obtained, slight modifications were made by way of replacement and redrawing of some items of the test. The number of items on the test was, however, kept unchanged.

ADMINISTRATION II

INTRODUCTION

The administration II was undertaken to conduct a rigorous item analysis. The main purpose of this analysis was to select comparatively valid items for inclusion in the final form of the test.

METHOD

Subjects.

A random sample of 278 (50% of the population) students of class IX drawn by means of a table of random numbers (Dixon et al., 1951, pp.290-294) from 11 High Schools (government and non-government) of Rajshahi Town in Bangladesh was used as subjects. This sample consisted of 66% male and 34% female students, of which 50% belonged to Humanities group and 50%, to Science group.

Tests used.

NLTMA consisting of 230 items.

Procedure.

The test was administered on a group and not individually. Desks were placed widely apart in a suitable room. Subjects were requested to sit one behind another. Care was taken to see that they felt at home and enjoyed the testing situation. Each of them was provided with a pencil and a rubber eraser. They were requested to fill up the blanks at the

top of the answer sheet. Then the test booklet was distributed to each of them. Instructions for taking the test were read out in Bengali, and in order to make these clear to the subjects, chalk and blackboard were utilized. Praise and encouragement were given during administration to induce good motivation. Subjects were asked to take all the time they required and not to be in a hurry. They were also asked not to go to the next page until they had completed the previous one.

There was one scoring key for each of the four subtests. The key was placed over the answer sheet in such a way that the markings could be seen through the holes. The keys were so made as to block responses other than right ones. The testee got one mark for each right answer and zero for wrong one. The summation of marks obtained in the subtests constituted the total score on the test.

RESULTS

The raw scores obtained by 278 pupils on the NLTMA were translated into a frequency distribution as shown in Table III. Fig.1 graphically represents the comparative distributions of raw, smoothed and expected frequencies. The graphs of raw and smoothed frequencies suggest that the distribution is negatively skewed. The skewness of the distribution was estimated by using the formula,

$$sk = \frac{3(M - Mdn)}{\sigma} \quad \ldots \quad (9)$$

(Garrett, 1960, p. 100). The skewness was found to be -0.345 ,which may

TABLE III

Showing frequency distribution of NLTMA scores.

Group x-values	Frequency	Smoothed frequency
200-209	0	1.67
190-199	5	5.33
180-189	11	10.66
170-179	16	20.66
160-189	35	28.66
150-159	35	33.66
140-149	31	34.00
130-139	36	32.33
120-129	30	30.00
110-119	24	26.00
100-109	24	19.00
90-99	9	14.66
80-89	11	8.33
70-79	5	6.00
60-69	2	2.66
50-59	1	1.33
40-49	1	1.00
30-39	1	0.66
20-29	0	0.66
10-19	1	0.33

N = 278

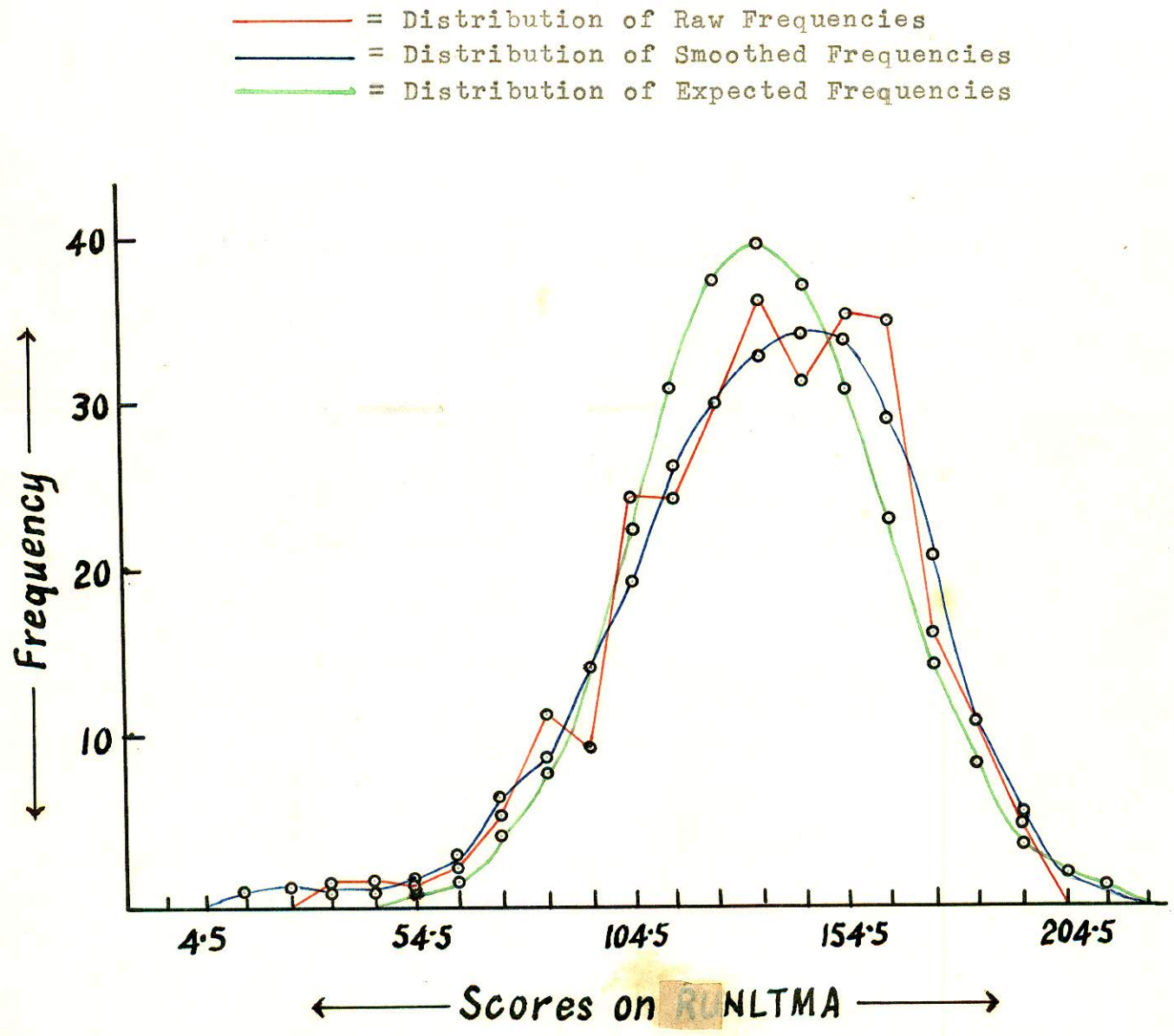


Fig. 1

Frequency distribution of NLTMA scores.

indicate that the distribution was slightly negatively skewed. The kurtosis was calculated by employing the formula,

$$ku = \frac{Q}{P_{90} - P_{10}} \quad \ldots \quad \ldots \quad (10)$$

(Garrett, 1960, p. 102) and found to be 0.284, which being greater than 0.263 (kurtosis for mesokurtic/normal distribution) by 0.021 may warrant that the distribution was slightly platykurtic.

In order to test if the sample was unbiased and representative, the frequency distribution was first graduated by comparison with a normal probability distribution and χ^2 - test was then applied. A trial hypothesis was framed up. The trial hypothesis was : The NLTMA distribution will not differ significantly from the normal distribution. The results have been shown in Table IV *.

In consultation with relevant table for probability integral of χ^2 - distribution, it was found that the χ^2 of 5.992 was significant at 0.81 level for 10 degrees of freedom. For 10 df, χ^2 should read 18.31 and 23.21 in order to be significant at 0.05 and 0.01 levels respectively (Pearson et al., 1966, pp. 129-130). The rejection of the null (trial)

---

* Table IV shows the comparison between the distributions of observed (smoothed) frequencies and expected frequencies, and that between the observed (raw) frequencies and expected frequencies has been reported at page 67 .

TABLE IV

NLTMA : Graduation by a normal curve and a test of goodness of fit.

	Group boundary x-values	$x - \mu$	$X = \frac{x-\mu}{\sigma_c}$	P(X)	$\Delta P(X)$	$N \Delta P(X)$	Observed frequency (smoothed)	$f_o - f_e$	$\frac{(f_o - f_e)^2}{f_e}$
	209.50	+73.70	+2.3500	0.99061	0.00099	0.275	1.670		
	199.50	+63.70	+2.3125	0.98962	0.03304	9.185	5.330	-2.460	0.6397
	189.50	+53.70	+1.7123	0.95658	0.03836	10.664	10.670	+0.006	0.000003
	179.50	+43.70	+1.3932	0.91822	0.05750	15.985	20.670	+4.685	1.3168
	169.50	+33.70	+1.0746	0.85872	0.08364	23.252	28.670	+5.418	1.2625
	159.50	+23.70	+0.7557	0.77508	0.10617	29.515	33.670	+4.155	0.5849
N = 278	149.50	+13.70	+0.4369	0.66891	0.12195	33.902	34.000	+0.098	0.0021
	139.50	+ 3.70	+0.1180	0.54696	0.12653	35.175	32.330	-2.845	0.2301
h = 10	129.50	- 6.30	-0.2008	0.42043	0.11879	33.023	30.000	-3.023	0.2767
μ = 135.80	119.50	-16.30	-0.5197	0.30164	0.09316	25.898	26.000	+0.102	0.0004
	109.50	-26.30	-0.8386	0.20848	0.08495	23.616	19.000	-4.616	0.9002
σ_c = 31.36	99.50	-36.30	-1.1575	0.12353	0.05361	14.903	14.670	-0.233	0.0037
(after	89.50	-46.30	-1.4764	0.06992	0.03362	9.346	8.330	-1.016	0.1104
Sheppard's	79.50	-56.30	-1.7953	0.03630	0.01905	5.296	6.000		
correction)	69.50	-66.30	-2.1142	0.01725	0.00976	2.713	2.670		
	59.50	-76.30	-2.4330	0.00749	0.00453	1.259	1.330		
df = 13-(2+1)	49.50	-86.30	-2.7519	0.00296	0.00189	1.525	1.000	+2.588	0.6643
= 10	39.50	-96.30	-3.0707	0.00107	0.00072	0.200	0.670		
	29.50	-106. 30	-3.3897	0.00035	0.000246	0.068	0.670		
	19.50	-116. 30	-3.7085	0.000104	0.000076	0.021	0.330		
	9.50	-126. 30	-4.0274	0.000028	-	-	-		

$\chi^2 = 5.991803$

hypothesis stands discredited on the basis of this statistical evidence. Hence the departure from normality cannot be established as significant on the basis of this test. The sample may, therefore, be regarded as unbiased and representative.

Item Analysis.

In view of the sample having been found to be unbiased and representative a rigorous item analysis was performed with the data. The results of analysis have been shown for each of the test-samples in Tables V, VI, VII and VIII. The item analysis was, however, carried out in the manner outlined in chapter III. The frequencies of right or wrong responses in accordance with the design shown in Table I (vide page 27) have not been presented , inasmuch as this is implied in the following tables.

TABLE V

Showing item analysis on Test A
(Figure Matching Test)

#	Difficulty Index (corrected for chance success) P_c	Item Variance pq	Validity Index r_b	Remarks
2	0.81	0.15	0.45	Rejected
3	0.59	0.24	0.29	Selected
4	0.47	0.25	0.42	,,
5	0.79	0.16	0.49	,,
7	0.55	0.24	0.41	,,
10	0.47	0.25	0.33	,,
11	0.77	0.17	0.64	,,
15	0.55	0.24	0.66	,,
20	0.81	0.15	0.59	Rejected
21	0.68	0.21	0.73	Selected
23	0.74	0.19	0.55	,,
24	0.76	0.18	0.61	,,
26	0.76	0.18	0.62	,,
28	0.80	0.16	0.57	,,
30	0.86	0.12	0.50	Rejected
31	0.69	0.21	0.45	Selected
32	0.55	0.24	0.55	,,
35	0.24	0.18	0.66	,,
36	0.48	0.25	0.58	,,
38	0.65	0.23	0.60	,,
41	0.80	0.16	0.52	,,
44	0.66	0.22	0.36	,,
49	0.72	0.20	0.50	,,

Number of items selected = 20

TABLE VI

Showing item analysis on Test B
(Sign Patterning Test)

#	Difficulty Index (corrected for chance success) P_c	Item variance Pq	Validity Index r_b	Remarks
1	0.57	0.24	0.54	Selected
2	0.81	0.15	0.59	Rejected
3	0.34	0.22	0.73	Selected
4	0.80	0.16	0.60	Rejected
5	0.23	0.17	0.43	Selected
6	0.56	0.24	0.60	,,
7	0.65	0.22	0.29	Rejected
8	0.65	0.22	0.70	Selected
11	0.76	0.18	0.66	,,
13	0.75	0.18	0.62	,,
14	0.67	0.22	0.60	,,
15	0.87	0.11	0.50	Rejected
17	0.47	0.25	0.64	Selected
18	0.35	0.22	0.68	,,
20	0.27	0.19	0.31	,,
22	0.76	0.18	0.46	Rejected
23	0.64	0.23	0.30	,,
25	0.72	0.20	0.52	Selected
26	0.72	0.20	0.29	Rejected
28	0.56	0.24	0.30	,,
31	0.60	0.24	0.38	Selected
32	0.32	0.21	0.60	,,
34	0.27	0.19	0.68	,,
36	0.26	0.19	0.28	,,
37	0.64	0.23	0.35	,,
39	0.50	0.25	0.42	,,
42	0.43	0.24	0.52	,,
43	0.46	0.25	0.49	,,
44	0.22	0.17	0.41	,,
45	0.41	0.24	0.46	,,
46	0.63	0.23	0.32	,,
47	0.40	0.24	0.32	,,
52	0.66	0.22	0.31	,,

Number of items selected = 25

TABLE VII

Showing item analysis on Test C

(Series Completion Test)

#	Difficulty Index (corrected for chance success) P_c	Item variance Pq	Validity Index r_b	Remarks
2	0.59	0.24	0.69	Rejected
5	0.58	0.24	0.71	Selected
6	0.46	0.24	0.77	,,
8	0.52	0.25	0.88	,,
9	0.55	0.24	0.78	,,
12	0.47	0.25	0.69	,,
13	0.53	0.25	0.81	,,
22	0.59	0.24	0.74	,,
23	0.56	0.24	0.84	,,
25	0.41	0.24	0.81	,,
26	0.44	0.24	0.82	,,
29	0.44	0.24	0.72	Rejected
36	0.60	0.24	0.80	Selected
37	0.54	0.24	0.76	,,
40	0.27	0.19	0.53	,,
41	0.50	0.25	0.79	,,
44	0.54	0.24	0.69	,,
46	0.58	0.24	0.82	,,
54	0.46	0.25	0.61	,,
61	0.50	0.25	0.86	,,
69	0.40	0.24	0.74	,,

Number of items selected = 20

TABLE VIII

Showing item analysis on Test D

(Pattern Synthesis Test)

#	Difficulty Index(corrected for chance success) P_c	Item variance Pq	Validity Index r_b	Remarks
1	0.31	0.21	0.28	Rejected
2	0.60	0.24	0.73	,,
6	0.78	0.17	0.65	,,
9	0.42	0.24	0.61	Selected
12	0.60	0.24	0.77	,,
13	0.61	0.23	0.65	,,
17	0.41	0.24	0.64	,,
18	0.51	0.25	0.66	,,
19	0.57	0.24	0.47	,,
24	0.27	0.19	0.45	,,
25	0.59	0.24	0.78	,,
27	0.52	0.25	0.77	,,
28	0.39	0.23	0.75	,,
29	0.48	0.25	0.40	,,
30	0.47	0.25	0.59	,,
36	0.45	0.24	0.63	,,
38	0.44	0.24	0.79	,,
39	0.58	0.24	0.79	,,
40	0.53	0.25	0.48	,,
41	0.61	0.23	0.64	Rejected
42	0.28	0.20	0.68	Selected
43	0.34	0.22	0.72	,,
44	0.38	0.25	0.64	Rejected
45	0.32	0.21	0.43	Selected
46	0.38	0.23	0.71	,,

Number of items selected = 20

DISCUSSION

Tables V,VI, VII and VIII contain the main features of item analysis and the rationale underlying the selection of items for inclusion in the test. In Test A, the difficulty indices of items (selected) range from 0.24 to 0.80, which may explain why item No. 2 (p_c = .81) and item No.20 (p_c = .81) have been rejected. Item variance ranges from 0.16 to 0.25 (0.25 being the maximum possible variance that an item can have) ; whereas the validity indices ranges from 0.29 to 0.73. Item Nos. 7, 15 and 32 have the same difficulty index (i.e., 0.55) but different validity indices (i.e., .41, .66 and .55 respectively). As a rule, item No.15 should be selected as it has the maximum validity index. If this were done, this would result in the reduction of the number of items on the subtest by 2. The ultimate result would be the lowering of the coefficient alpha, which the test is not designed to entertain. Moreover, the difficulty indices of these items fall within the optimum range. So, these items were included in the test, their difficulty indices in the lower 27% group being of help in arranging them in order. Thus, item Nos. 7,32 and 15 occupy the positions indicated by the serials 14, 15 and 16 respectively in the final form. Similar arguments may apply to the selection of item Nos. 41 and 28 (p_c = .80) as well as 24 and 26 (p_c = .76). They occupy the positions indicated by the serial Nos. 1,2,5 and 6 respectively in the final form. The same arguments may also hold good for the selection of item Nos.4 and 10.

These items are however, indicated by the serial Nos. 19 and 18 respectively on the final test. The number of items selected was 20.

In Test B, item No. 8 has been selected in preference to item No.7, though both have the same difficulty index (0.65). The magnificent difference in their validity indices (0.29 for No. 7 and 0.70 for No. 8) may explain the reason for doing so. In this subtest the difficulty indices of items(selected) range from 0.22 to 0.76, which may explain why certain items having greater p_c's (hence, easier) and quite satisfactory validity indices have been rejected. The item variances range from 0.17 to 0.25 and the validity indices, from 0.28 to 0.73. The number of items selected was 25.

In Test C, the difficulty indices range from 0.27 to 0.60. The item variances range from 0.19 to 0.25 and the validity indices, from 0.53 to 0.88. Though item Nos. 2 and 22 have the same difficulty index (0.59), the item No. 22 has been selected in view of its higher validity index (0.74). Similar arguments may apply to item Nos.26 and 29, where the former has been selected. The number of items selected was 20.

In Test D, the validity indices range from 0.40 to 0.79. The difficulty indices range from 0.27 to 0.61 and the item variance, from 0.19 to 0.25. Item No.12 has been selected in preference to item No. 2 in view of the greater validity index which the former possesses (i.e., 0.77). Item No. 24 has been selected in preference to item No.1, as the former has a higher validity index (0.45). The number of items thus selected was 20.

Hence the total number of items on the test reached 85 down from a lot of 230 items. The coefficient alphas calculated by using formula 2(vide page 28)

were found to be .74, .54, .71 and .72 respectively for Test A, Test B, Test C and Test D. The coefficient alpha for the whole test was found to be .96. This means that the present test may be homogeneous and that it may yield unambiguous scores.

ADMINISTRATION III

INTRODUCTION

The administration III was conducted on the standardization sample which was statistically tested for representativeness and umbiasedness. The main objectives of this study were to estimate the reliability (split-half) and the validity (predictive and factorial) of the test, to prepare the norm and to show the comparison between the performances of boys and girls, of students of Humanities and Science groups, and of students of government and non-government schools on the present test.

METHOD

Subjects.

A random sample of 270 (25% of the population) students of class IX drawn by means of a table of random numbers (Dixon et al., 1951, pp.290-294) from 13 High Schools (both government and non-government) of Rajshahi Town in Bangladesh was used as subjects. This sample comprised 63.33% male and 36.67% female students, of which 58.52% belonged to Humanities group and 41.48%, to Science group.

Tests used.

NLTMA (final form) consisting of 85 items.

Procedure.

The test was administered on a group. Desks were placed widely apart in a room suited to the purpose. Subjects were requested to sit one behind another. Care was taken to see that they felt at home and enjoyed the testing situation. Each of them was provided with a pencil and a rubber eraser. They were requested to fill up the blanks at the top of the answer sheet. Then the test booklet was distributed to each of them. Instructions for taking the test were read out in Bengali, and, in order to make these clear to the subjects, practice items were drawn on the blackboard and the working out of their solutions, explained. Praise and encouragment were given during administration to get them motivated. Subjects were asked to take all the time they needed and not to be in a hurry. They were also asked not to go to the next page until they had completed the previous one.

There was one scoring key for each of the four subtests. The key was placed over the answer sheet in such a way that the markings could be seen through the holes. The keys were so made as to block responses other than the right ones. The testee got one mark for each right answer and zero for wrong one. The summation of marks obtained in the subtests constituted the total score on the test.

RESULTS

The raw scores obtained by 270 pupils on the test were translated into a frequency distribution (vide Table IX). Fig.2 graphically represents the comparative distributions of raw, smoothed and expected frequencies. The

skewness of the distribution was determined by using formula 9 (vide page 36). The skewness was found to be -.400, which may warrant that the distribution was slightly negatively skewed. The kurtosis was calculated by using formula 10 (vide page 39), and was found to be .253, which, being less than .263 (kurtosis for mesokurtic/normal distribution) by .01, may indicate that the distribution was very slightly leptokurtic.

TABLE IX

Showing frequency distribution
of NLTMA (final form) scores.

Group x-values	Frequency	Smoothed frequency
82 and above	0	1.33
77-81	2	6.33
72-76	17	18.00
67-71	35	33.33
62-66	48	43.66
57-61	48	45.00
52-56	39	39.33
47-51	31	28.66
42-46	16	20.66
37-41	15	14.00
32-36	11	11.33
27-31	8	6.33
26 and below	0	2.66
	N = 270	

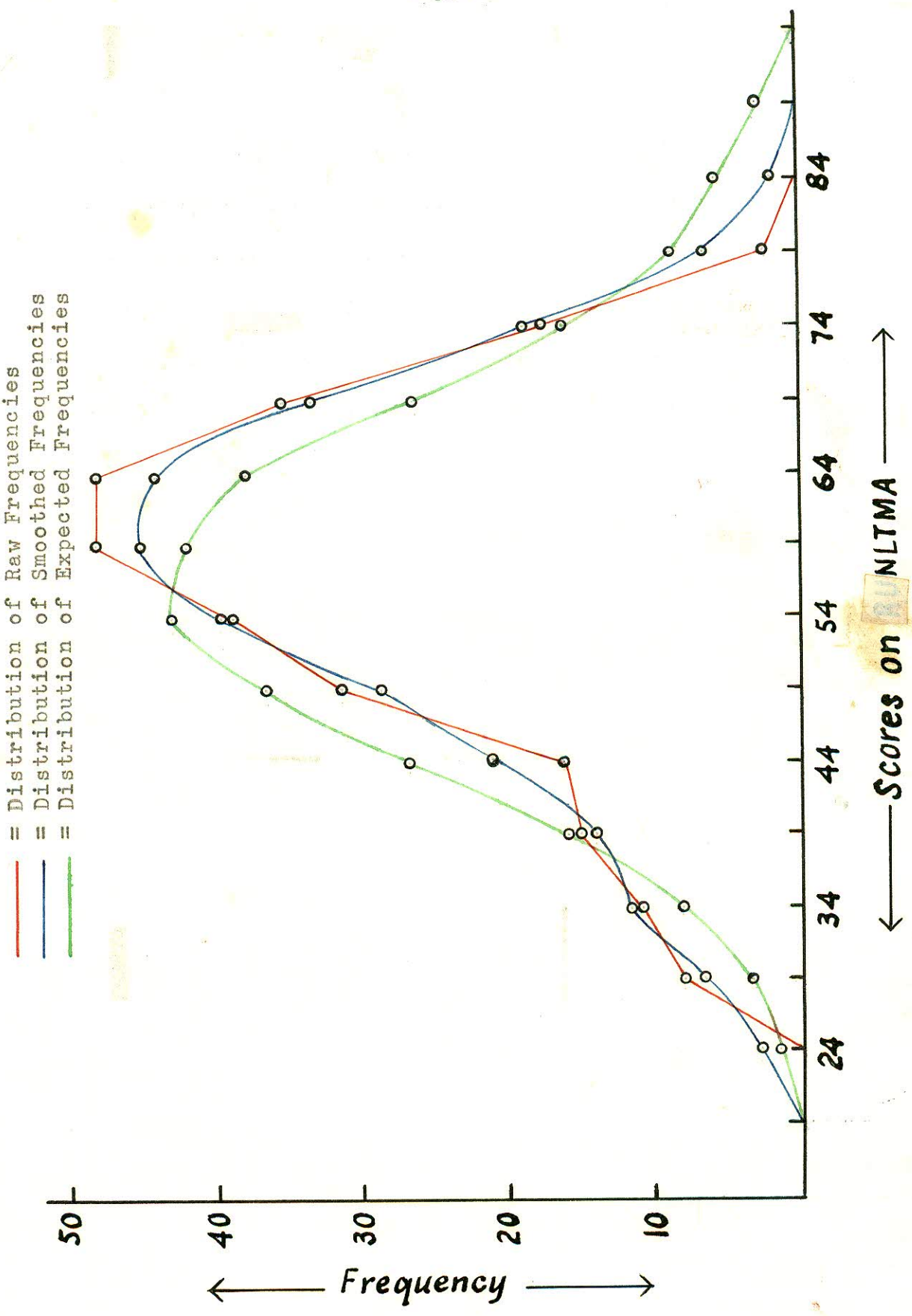


Fig.2

Frequency distribution of NLTMA(final form) scores.

T-78

In order to test whether the sample was unbiased and representative, the frequency (smoothed) distribution was first graduated in reference to the normal probability distribution and χ^2-test was then applied.

A trial (null) hypothesis was then framed up. The trial hypothesis was : The NLTMA distribution will not differ significantly from the normal distribution. The results have been shown in Table X.

In consultation with the relevant table for probability integral of χ^2-distribution, it was found that the χ^2 of 11.73 was significant at 0.23 level for 9 degrees of freedom. For 9 df, χ^2 should read 16.90 and 21.20 in order to be significant at 0.05 and 0.01 levels respectively (Pearson et al., 1966, pp. 129-130). The rejection of the null hypothesis, therefore, stands unwarranted on the basis of this statistical evidence. As such, the departure from normality cannot be established as significant on the basis of this test. The sample may, therefore, be regarded as unbiased and representative.

Tests for a power test.

Results showed that, since u-scores were zero, $M_w = M_x = 28.6$, $\sigma_w = \sigma_x = 11.55$, and that $\frac{\sigma_u}{\sigma_x} = \frac{0}{11.55} = 0$. Hence, the criteria for a power test may be taken as satisfied.

Reliability.

The reliability of the test was calculated by using formulae 5 and 6 (vide page 30) , and was found to be .81 (significant beyond .01 level for 268 degrees of freedom).

TABLE X*

NLTMA (final form) : Graduation by a normal curve and a test of goodness.

	Group boundary x-values	$x-\mu$	$X = \frac{x-\mu}{\sigma_c}$	P(X)	$\Delta P(X)$	$N.\Delta P(X)$	Observed frequency (smoothed)	$\frac{f_o^2}{f_e}$
	82 and above			1.0000	0.0197	5.32	1.33	0.33
	81.50	+24.95	+2.06028	0.9803	0.0304	8.21	6.33	4.99
	76.50	+19.95	+1.64740	0.9496	0.0588	15.88	18.00	20.40
N = 270	71.50	+14.95	+1.23452	0.8908	0.0968	26.16	33.33	42.46
μ = 56.50	66.50	+9. 95	+0.82163	0.7940	0.1395	37.66	43.66	50.62
h = 5	61.50	+4. 95	+0.40875	0.6555	0.1555	41.98	45.00	48.23
σ_c = 12.11 (After Sheppard's adjustment)	56.50	0. 00	0.00000	0.5000	0.1592	42.98	39.33	33.33
	51.50	-5. 05	-0.41701	0.3408	0.1348	36.40	28.66	22.56
df = 12-(2+1) = 9	46.50	-10.05	-0.82989	0.2060	0.0986	26.62	20.66	16.04
	41.50	-15.05	-1.24277	0.1074	0.0580	15.66	14.00	12.52
	36.50	-20.05	-1.65566	0.0494	0.0298	8. 05	11.33	14.70
	31.50	-25.05	-2.06854	0.0196	0.0129	3.48	6.33	15.45
	26 and below	-30.05	-2.48142	0.0065	0.0065	1.75	2.66	

$\therefore \chi^2 = \sum \frac{f_o^2}{f_e} - N = 281.73 - 270 = 11.73$

* Table X shows the comparison between the distributions of smoothed frequency and expected frequency. This has been done because the distribution of the smoothed frequency better fits the normal distribution.

The index of reliability coefficient was given by $r_{t\infty} = \sqrt{.81} = .90$, which was much closer to the coefficient alpha (.96). The test-retest reliability coefficient was determined by applying the test again to the same sample after 2 months. This was found to be .68 (significant beyond .01 level, i.e., $P < .01$).

The significance of the difference between the initial test scores and retest scores was tested through Student t-analysis. The observed difference was .60 (i.e., 59.9-59.3) and t was found to be .34. A consultation with the relevant table of t-distribution showed that it should read 1.66 and 1.99 in order to be significant at .10 and .05 levels respectively for 87 degrees of freedom. The obtained t of .34 was far short of these values and hence was not significant ($P > .10$).

Validity.

The predictive validity of the test was found by estimating the PM correlation between the test scores and the marks obtained by the same students in the School Annual Examination which was held a few days after the third administration of the test. The correlation was found to be .45 (significant beyond .01 level, i.e., $P < .01$).

The factorial validity of the test was determined by the complete centroid method of Professor L.L.Thurstone (Guilford, 1954,pp.470-524; Thurstone, 1947,pp.161-170). The results have been shown in Tables XI to XIX.

TABLE XI

Showing correlation matrix and extraction of the first centroid factor.*

Test		D	C	B	A	Check sum
1	D	-	.53	.33	.30	1.16
2	C	.53	-	.36	.15	1.04
3	B	.33	.36	-	.40	1.09
4	A	.30	.15	.40	-	0.85
	S_1	1.16	1.04	1.09	0.85	4.14
	D	0.53	0.53	0.40	0.40	1.86
$S_1 + D = E$		1.69	1.57	1.49	1.25	6.00 = T
$mE = a_1$		0.689	0.641	0.608	0.510	2.448

$$\sqrt{T} = 2.4495$$

$$\sqrt{\frac{1}{T}} = 0.40825 = m$$

$$\therefore \quad mT = 2.4495$$

* The data processing was done manually. Had there been a computer, the job would have been much easier.

TABLE XII

Showing computation of the first factor residuals and extraction of the second centroid factor.

k_1		-.680	-.641	-.608	-.510	$-2.448 = -v_1$	
a_1	Test	1	2	3	4	Check sum	
.689	1		.088	-.095	-.051	-.058	
.641	2	.088		-.029	-.177	-.118	
.608	3	-.095	-.029		.090	-.034	
.510	4	-.051	-.177	.090		-.138	
$2.448 = v_1$						-.348	
	S_1	1.160	1.040	1.090	0.850	4.140	
	$k_1 v_1$	-1.690	-1.570	-1.488	-1.248	-5.966	$(-v_1)(v_1) = -5.993$
	k_1^2	0.475	0.411	0.369	0.260	1.515	$\sum a_1^2 = 1.515$
$S_1+k_1v_1+k_1^2$		-.055	-.119	-.029	-.138	-.341	
	S_2	-.058	-.118	-.034	-.138	-.348	
$-S_2/2=A$		.029	.059	.017	.069	.174	
+4		-.022	-.118	.107	(.069)	.036	
B+3		-.117	-.147	(.107)	(.159)	.002	
-2B=C		.234	.294	-.214	-.318	-.004	$\sum \lvert C \rvert = 1.060$
D		.095	.177	-.095	-.177		$\sum \lvert D \rvert = 0.544$
C+D=E		.329	.471	-.309	-.495	1.604 = T	1.604
mE=a_2		.259	.372	-.244	-.391	1.266	

$$\sqrt{T} = 1.2665$$

$$\sqrt{\frac{1}{T}} = 0.78958 = m$$

$$\therefore mT = 1.2665$$

TABLE XIII

Showing computation of the second factor residuals.

k_2		-.259	-.372	.244	.391	.004 $=-v_2$	
a_2	Test	1	2	3	4	Check sum	
.259	1		-.008	-.031	.050	.011	
.372	2	-.008		.062	-.038	.016	
-.244	3	-.032	.062		-.005	.026	
-.391	4	.050	-.038	-.005		.007	
-.004 = v_2						.060	
	S_2	-.058	-.118	-.034	-.138	-.348	
	k_2v_2	.001	.001	-.001	-.002	-.001	$(-v_2)(v_2)$=-.00002
	k_2^2	.067	.138	.059	.153	.417	$\sum a_2^2$ = .417
$S_2+k_2v_2+k_2^2$		.010	.021	.024	.013	.068	
	S_3	.011	.016	.026	.007	.060	
$-S_3/2$ = A		-.005	-.008	-.013	-.003		

TABLE XIV

Centroid Factor Matrix

Test	Factor		Factor variances		Communalities	
	a_1	a_2	a_1^2	a_2^2	$h^2_{obt.}$	h^2_{guess}
1	.689	.259	.475	.067	.542	.53
2	.641	.372	.411	.138	.549	.53
3	.608	-.244	.370	.059	.429	.40
4	.510	-.391	.260	.153	.413	.40
	Σa_k^2		1.516	0.417	1.933	
			78.4%	21.6%	100.00%	

Orthogonal Rotation of Reference Axes.

In order to identify and interpret psychologically the centroid factors, these were put to orthogonal rotation. The orthogonal rotation has been shown in Fig.3. The rotation was carried on until a good one was found (Guilford, 1954, p. 485 ; Thurstone, 1947, pp. 341-343). The results have been shown in Tables XV to XIX.

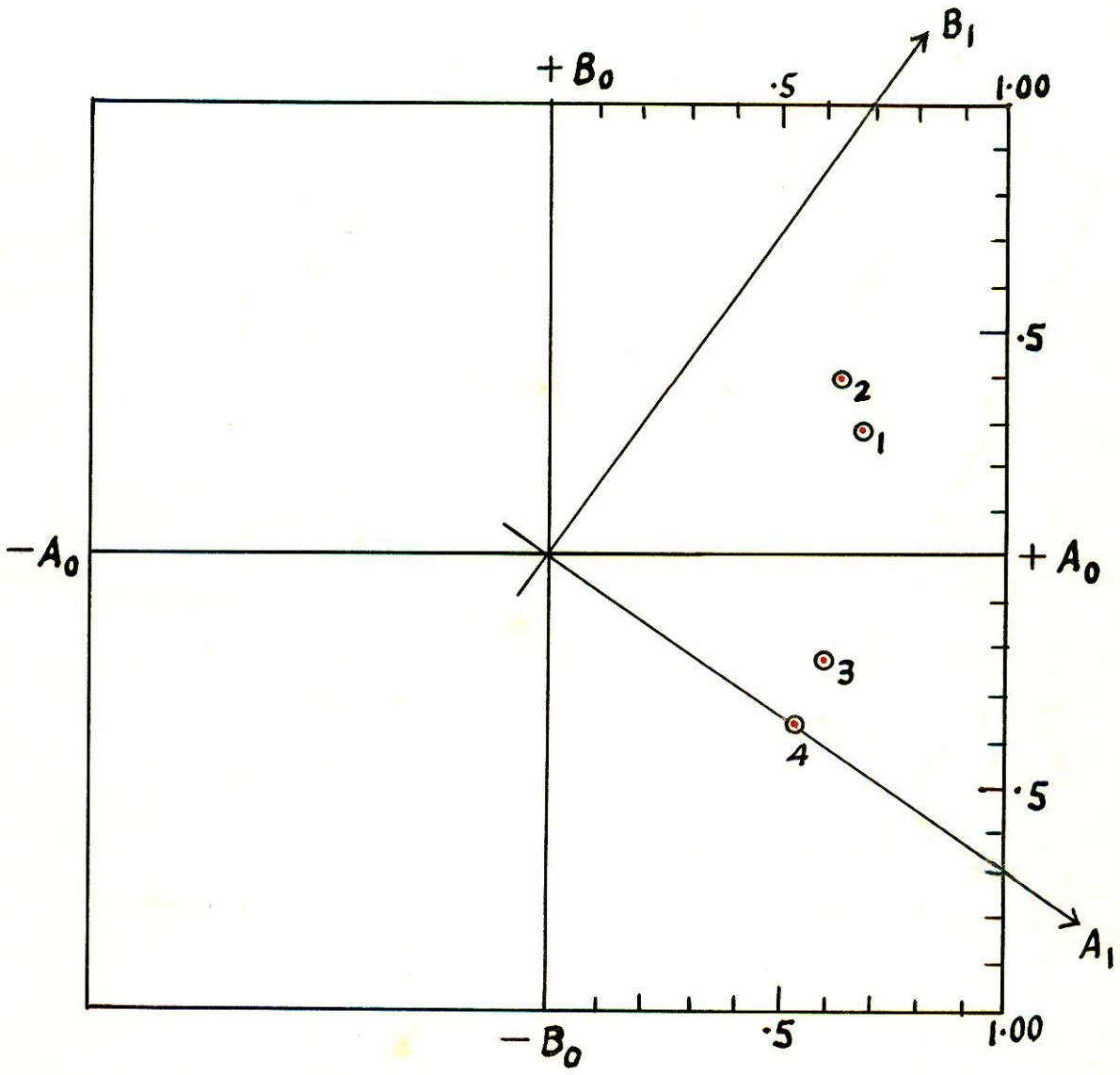


Fig.3
Orthogonal Rotation of Reference Axes

TABLE XV

Centroid Matrix (C)

Test	A_o	B_o
1	.689	.259
2	.641	.372
3	.608	-.244
4	.510	-.391

TABLE XVI

Direction Numbers(L_{o1})

	A_1	B_1
A_o	1.00	.69
B_o	- .69	1.00
$\sum l^2$	1.4761	1.4761
$\sqrt{\sum l^2}$	1.2149	1.2149
D_{o1}	0.8231	0.8231

TABLE XVII

Transformation Matrix (T_{01}) (Direction Cosines)

	A_1	B_1
A_0	.823	.568
B_0	-.568	.823

TABLE XVIII

Rotated Factor Matrix (O_1)

	Test	A_1	B_1
$C.T_{01}$ =	1	.420	.604
	2	.316	.670
	3	.639	.144
	4	.642	-.032

TABLE XIX

Showing factor matrix and communalities after orthogonal rotation of reference axes.

Test		Factors		Factor variances		Communalities	
		a_1	a_2	a_1^2	a_2^2	$h^2_{obt.}$	$h^2_{guess.}$
1	D	.420	.604	.176	.365	.541	.53
2	C	.316	.670	.099	.449	.548	.53
3	B	.639	.144	.408	.011	.419	.40
4	A	.642	-.032	.412	.001	.413	.40
		Σa_k^2		1.095	0.826	1.921	
				57.00%	43.00%	100.00%	

Significance of factor loadings.

The significance of the factor loadings on the subtests of the NLTMA was statistically tested by calculating their standard errors by using formula 8 (vide page 31). The SE's of the loadings of factor a_1 on Test A (Figure matching test), Test B (Sign patterning test), Test C (Series completion test) and Test D (Pattern synthesis test) were respectively found to be .041, .042, .063, and .058. These loadings may be considered to be significant, since the multiplication of each SE by 2 yielded an amount less than the corresponding factor loading. Again, the SE's of the loadings of factor a_2 on Test A, Test B, Test C and Test D were estimated to be .069, .064, .039, and .044 respectively. Multiplication of each SE by 2 yielded an amount less than the corresponding factor loading. The loadings of factor a_2 may, therefore, be regarded as statistically significant, excepting that on Test A.

DISCUSSION

Tables XIV and XIX show that the obtained communalities in the centroid factor matrix and the rotated factor matrix are almost the same. It is also seen from Table XIX (vide page 61) that the first factor has taken 57% and the second factor has taken 43% of the total variance.

Factor a_1 is rich on the subtests 3 and 4 (Sign pattering test and Figure matching test respectively). Subtests 1 and 2 (Pattern synthesis test and Series completion test respectively) are loaded with factor a_1.

The lowest loading of factor a_1 is on the subtest 2 (Test C). This may be the well-known "Perceptual speed" factor (P), involving the seeing of spatial and perceptual relations with rapidity and accuracy. Factor a_1 may be called general factor by virtue of its having +ve loadings on all the subtests (Thurstone, 1947, p.121).

Factor a_2 is rich on subtests 1 and 2 (Pattern synthesis test and Series completion test respectively). Test C (Series completion test) has a loading of .670 with factor a_2 , and Test D (Pattern synthesis test) ,a loading of .604 with factor a_2 . Factor a_2 may be taken as the "Spatial visualization" factor (S). The subject has to manipulate objects imaginally in space,in addition to seeing relations,in solving the problems in these subtests. The lowest +ve loading of .144 is on Test B, which may mean that in Sign patterning test (Test B) the subject requires but a little spatial visualization to solve problems. The negative loading (-.032) of factor a_2 on Test A (Figure matching test) may indicate that this factor is negatively related to factor a_1 on this test. Factor a_2 may be called bipolar inasmuch as it has both +ve and -ve loadings (Thurstone, 1947,p.121).

The factorial validities of Series completion test (Test C) as a measure of spatial visualization factor (factor a_2) and perceptual speed factor (factor a_1) are respectively .670 and .319. The factorial validities of Pattern synthesis test (Test D) as a measure of perceptual speed factor and spatial visualization factor are .420 and .604 respectively.The factorial validities of Figure matching test (Test A) as a measure of perceptual speed

factor and spatial visualization factor are .642 and -.032 respectively. The factorial validities of Sign patterning test (Test B) as a measure of perceptual speed factor and spatial visualization factor are .639 and .144 respectively. As such, Test C (Series completion test) and Test D (Pattern synthesis test) may be regarded as valid measures of spatial ability, and Test A (Figure matching test) and Test B (Sign patterning test), as valid measures of the ability to perceive spatial relations with rapidity and accuracy.

Development of R-scale for preparing the norm.

The R-scale has been named after 'Rajshahi University Test of Mental Ability' (RUTMA). The experimental data from which the scale has been developed have been obtained from the administration of the following tests : (1) The Rajshahi University Test of Mental Ability (RUTMA) and (2) the Non-Language Test of Mental Ability (NLTMA).

The RUTMA is a test of general intelligence (verbal) for adults and the NLTMA is a non-verbal test of intelligence for secondary school students (Class IX-X). These were administered to samples of 198 and 270 subjects respectively. The sample of 198 consisted of 50% science and 50% arts students (Postgraduate previous)of Rajshahi University in Bangladesh, of which 15% were female and 85%,male .Detailed information regarding the sample used for NLTMA has already been presented (vide page 48).

The results obtained from their administration were translated into frequency distributions. These distributions were then graduated in reference to the normal probability curve and the goodness of fit was tested through χ^2-analysis. The skewness and kurtosis were calculated by employing formulae 9 and 10 (vide pages 36 and 39) respectively. The results of RUTMA have been presented in Table XX. The results of NLTMA have been shown in Table X (vide page 52a).

In consultation with the relevant table for the probability integral of χ^2-distribution we found the χ^2 of 8.773 (vide Table XX) to be significant at 0.64 level for 11 degrees of freedom. For 11 df, χ^2 should read 22.5 and 19.7 in order to be significant at 0.01 and 0.05 levels(Pearson et al., 1966, pp.129-130) respectively. Therefore, χ^2 of 8.773 was not significant at 0.01 and 0.05 levels. The rejection of the null hypothesis that the frequency distribution of scores on RUTMA will not differ significantly from the normal probability distribution stands unwarranted on the basis of this statistical evidence. As such, the departure from normality cannot be established as significant on the basis of this test. The sample may, therefore, be taken as unbiased and representative.

The skewness, was found to be -0.1325, which may indicate that the distribution was slightly negatively skewed. The kurtosis was found to be 0.278, which was greater than 0.263 (kurtosis for mesokurtic distribution) by 0.015 only. This statistical evidence may warrant that the distribution was slightly platykurtic.

TABLE XX

RUTMA : Graduation by a normal curve and a test of goodness of fit.

	Ground boundary x-values	$x-\mu$	$X=\frac{x-\mu}{\sigma_c}$	P (X)	$\Delta P(X)$	$N\Delta P(X)$	Observed frequency	$\frac{f_o^2}{f_e}$
	114.50	+45.55	+2.7292	0.99682	0.00438	0.87	3	
	109.50	+40.55	+2.4295	0.99244	0.00903	1.79	2	5.970
	104.50	+35.55	+2.1300	0.98341	0.01701	3.37	1	
	99.50	+30.55	+1.8304	0.96640	0.02930	5.80	6	6.207
	94.50	+25.55	+1.5308	0.93710	0.04620	9.15	13	18.470
	89.50	+20.55	+1.2313	0.89090	0.06665	13.20	15	17.045
	84.50	+15.55	+0.9317	0.82425	0.08791	17.41	14	11.523
N = 198	79.50	+10.55	+0.6321	0.73634	0.10610	21.01	20	19.038
h = 5	74.50	+ 5.55	+0.3325	0.63024	0.11712	23.19	26	29.150
μ = 68.95	69.50	+ 0.55	+0.0329	0.51312	0.11822	23.41	19	15.421
σ_c = 16.69	64.50	- 4.45	-0.2666	0.39490	0.10927	21.63	15	10.402
(after Sheppard's adjustment)	59.50	- 9.45	-0.5662	0.28563	0.09233	18.28	22	26.477
df = 14-(2+1)	54.50	-14.45	-0.8658	0.19330	0.07040	13.94	16	18.364
= 11	49.50	-19.45	-1.1660	0.12290	0.05142	10.18	11	11.886
	44.50	-24.45	-1.4649	0.07148	0.03265	6.46	7	7.585
	39.50	-29.45	-1.7645	0.03883	0.01933	3.83	2	
	34.50	-34.45	-2.0641	0.01950	0.01045	2.07	4	9.235
	29.50	-39.45	-2.3636	0.00905	0.00518	1.03	2	
	24.50	-44.45	-2.6633	0.00387				

$$\therefore \chi^2 = \sum \frac{f_o^2}{f_e} - N = 206.773 - 198.000 = 8.773$$

By actual counting, 66.16% of cases were found to lie between $M \pm 1\sigma$, 94.44% of cases, between $M \pm 2\sigma$ and 99.49% of cases, between $M \pm 2.5\sigma$ in the case of distribution of RUTMA scores. The remaining 0.51% of cases were covered up by $M \pm 3\sigma$. By following the same technique, 70.37% of cases were found to lie between $M \pm 1\sigma$, and 97.37% of cases, between $M \pm 2\sigma$ in the case of the distribution of NLTMA scores. The remaining 2.63% of cases were covered up by $M \pm 3\sigma$.

In another study with NLTMA on a random sample of 278 secondary school students (66% boys, 34% girls, 50% Humanities and 50% Science groups ; $\chi^2 = 2.914$, being significant at 0.98 level and not significant at 0.05 and 0.01 levels for 10 df) , 63.66% of cases were found to lie between $M \pm 1\sigma$, 96.76% of cases, between $M \pm 2\sigma$, 98.92% of cases, between $M \pm 3\sigma$, and 99.64% of cases, between $M \pm 4\sigma$. The remaining 0.36% of cases were taken care of by $M \pm 4.5\sigma$. The distribution of NLTMA scores in this case was found to be slightly negatively skewed (sk = -0.345), and slightly platykurtic (ku = 0.284, which > 0.263 by 0.021).

Now that the experimental evidences have been presented, it seems to be worthwhile to enter upon a brief discussion about some of the important scales in use. The assumption of equal percentile units on which the percentile scale is based holds strictly only when the distribution of scores is rectangular in shape; it does not hold true when the distribution is approximately normal. Distributions of raw scores are rarely , if ever , rectangular in form (Garrett, 1960, pp. 321-322). Again, an IQ is a ratio of

MA to CA, multiplied by 100. It changes with the change in the MA, provided the CA remains constant. But psychologists are not unanimous in respect of the magnitude of CA used in calculating an IQ. According to Binet it is 15 years and to Otis, 18 years. It is, however, recognized by psychologists that the distribution of intelligence as measured by intelligence tests approximately conforms to the normal probability curve. The σ-scale leaves half of the scores as - ve and half as + ve below and above the mean respectively. Moreover, σ-scores are small decimal fractions and hence somewhat awkward to treat of in calculation (Garrett, 1960, p.312; Guilford, 1966, p. 516). The σ-scale suffers from using too large a unit like the standard deviation itself. The T-scale ranges from -5σ to $+5\sigma$ with one-tenth of the SD as the unit of measurement (McCall, 1939, pp.505-508). This scale suffers from over refinement of the unit. In the measurement of higher mental functions, as Binet has commented, the greatest precision of measurement, though desirable, is not as essential as in measuring the simpler processes, because of the fact that individuals differ much more markedly in the former (Freeman, 1960, p. 100). Besides, the distribution of scores is not found, in practice, to range from -5σ to $+5\sigma$. The objection to the use of the range (-5σ to $+5\sigma$) is substantiated by a consideration of the results referred to above.

Considering what has been discussed above, the R-scale has been developed on the basis of normality of distribution of intelligence test scores. The

purpose was to prepare the norm for the present test. This scale ranges from -4.5 σ to +4.5 σ . The whole area between the curve and its asymptote is divided into 9 equal parts. The zero of the R-scale is located at -4.5 σ instead of at -5 σ of the T-scale. The raw scores are transformed into a normal distribution with a mean of 50 and an SD of 11 by virtue of formula 12. The unit of measurement is .2 σ in lieu of .1 σ of the T-scale. The steps followed in deriving the R-score have been shown below :

1. The class intervals were listed as usual.
2. The exact upper limits of the class intervals were listed.
3. The class frequencies were noted down.
4. The cumulative frequencies were listed.
5. The cumulative percentages were calculated.
6. The mean, M_0, was estimated.
7. The standard deviation, σ_0 , was calculated.
8. The observed raw scores were transformed into a normal distribution with a mean of 50 and an SD of 11 by virtue of formula 12.
9. The SD of 11 was multiplied by 0.2 .
10. The newly obtained scores (X_k) were graduated in terms of the figure, .2σ, to transform them into R-scores.

A general formula was developed for transforming the raw scores into a distribution with a desired mean and a standard deviation. The development of this formula has been presented below.

If X_0 be the observed score with M_0, the observed mean and σ_0, the

observed standard deviation and X be the desired score with M and σ as the desired mean and standard deviation respectively, then, on an assumption of normal distribution of scores in either case, we obtain,

$$\frac{X - M}{\sigma} = \frac{X_0 - M_0}{\sigma_0} \qquad \ldots\ldots \qquad (11)$$

By cross-multiplying and transposing, we have from (11),

$$(X - M)\,\sigma_0 = (X_0 - M_0)\,\sigma$$

$$\text{or,} \quad X\sigma_0 - M\sigma_0 = X_0\sigma - M_0\sigma$$

$$\text{or,} \quad X\sigma_0 = X_0\sigma - M_0\sigma + M\sigma_0$$

Dividing by σ_0, we have,

$$X = \frac{\sigma}{\sigma_0} X_0 - \frac{\sigma}{\sigma_0} M_0 + M$$

$$= \frac{\sigma}{\sigma_0} (X_0 - M_0) + M \qquad \ldots\ldots \qquad (12)$$

TABLE XXI *

Showing norms in terms of R-score.

(Boys and Girls together, Class IX, N = 270)

Score	Cumulative percentage	R-score
81	100	33
76	99.3	32
71	92.9	29
66	80.0	27
61	62.2	25
56	44.4	22
51	30.0	20
46	18.5	18
41	12.6	16
36	7.0	14
31	3.0	12

* The R-scores have been rounded off to the nearest whole numbers. The R-scores corresponding to any X-value falling within two upper limits can be calculated by interpolation.

TABLE XXII*

Showing norms in terms of R-score.

(Boys, Class IX, N = 171)

Score	Cumulative percentage	R-score
79	100	32
74	96.5	30
69	83.5	27
64	66.0	25
59	47.9	23
54	29.8	21
49	19.9	19
44	13.4	16
39	10.5	14
34	7.1	11

*The R-scores have been rounded off to the nearest whole numbers. The R-scores corresponding to any X-value falling within two limits can be obtained by interpolation.

TABLE XXIII *

Showing norms in terms of R-scores.

(Girls, Class IX, N = 99)

Score	Cumulative percentage	R-score
79	100	34
69	98.9	32
64	82.8	27
59	69.6	25
54	49.5	23
49	32.4	21
44	22.2	18
39	11.1	16
34	5.0	14
29	2.0	11

* R-scores have been rounded off to the nearest whole numbers. The R-scores corresponding to any X-values falling within two limits can be obtained by interpolation.

TABLE XXIV

Showing comparison between the performances of boys and girls on MLTMA.

	Mean	SD	N	CR
Boys(Class IX)	58.25	11.40	171	3.19
Girls (,, ,,)	53.75	11.10	99	

Since, for large samples, the distribution of CR's is known to be normal around the true difference of population means, in consultation with the table of normal probability distribution, we find that 99.8% of cases lie between $\pm 3.19\,\sigma_D$ and that 0.2% of cases fall beyond the limit. So, under the null hypothesis, we can expect CR's as large as or larger than $\pm$ 3.19 to occur by chance only 0.2 times out of 100 comparisons of means of samples of boys and girls of class IX on the test. A mean difference of $\pm$ 3.19 (a CR of $\pm$ 3.19) cannot arise as a sampling fluctuation from zero and is, therefore, significant ($P < .01$). As such, the null hypothesis that boys and girls will not significantly differ in their performance on the test is statistically discredited. In general, boys have, therefore, every likelihood of doing significantly better than girls on this test.

In order to test whether there is an equal variability among boys and girls, the SE of the difference between the SD's was determined by using the formula,

$$SE_{D_\sigma} = \sqrt{SE^2_{\sigma_1} + SE^2_{\sigma_2}} \quad \ldots \quad \ldots \quad (13)$$

(Garrett, 1960, pp.232-233). The obtained difference between the SD's was .30 and the CR was found to be .30. This CR of .30 was far short of its value at .05 level (i.e., 1.96) for 268 df.The obtained difference was , therefore,not significant (P>.05). So, the rejection of the null hypothesis stands discredited on the basis of this test. It, therefore, seems reasonable to think that there is equal variability among boys and girls in respect of their performance on this test.

TABLE XXV

Showing comparison between the performances of Science and Humanities students on NLTMA.

	Mean	SD	N	CR
Science students (class IX)	59.95	11.15	112	4.06
Humanities students (class IX)	54.30	11.40	158	

Since, for large samples, the distribution of CR's is known to be normal about the true difference between the population means, consulting the table of normal probability distribution, it is found that 99.98% of cases lie between $\pm$ 4.06σ_D and only .02% lie outside. So, under the null hypothesis, it is expected that a CR as large as or larger than $\pm$ 4.06 might occur by chance only .02 times in 100 comparisons of means of samples of students of Science and Humanities group (class IX) on NLTMA. A mean difference of $\pm$ 4.06 cannot arise as a sampling fluctuation from zero and is, therefore, significant ($P < .01$). As such, the null hypothesis that Science students will not differ significantly from Humanities students in their performance on the test cannot be established as statistically true on the basis of this test. In general, therefore, Science students may do significantly better than Humanities students on the test.

In order to test if there is an equal variability among Science and Humanities students, the SE of the difference between the SD's was calculated by using formula 13. The obtained difference between the SD's was .25 and the CR was found to be .25. This CR of .25 was far short of its value at .05 level for 268 df. The obtained difference was, therefore, not significant ($P > .05$). So, the rejection of the null hypothesis stands unwarranted on the basis of this test. It, therefore, seems reasonable to conclude that there is equal variability among Science and Humanities students with respect to their performance on the test.

TABLE XXVI

Showing comparison between the performances of government school students and non-government school students (class IX) on NLTMA

	Mean	SD	N	CR
Government school students	62.30	10.35	107	7.07
Non-government school students	52.90	11.05	163	

Since the CR is much higher , the difference between the mean performances of students of government schools and those of non-government schools may be taken as significant ($P < .01$). In general, therefore, the students of government schools who mostly come from the upper socio-economic stratum of the society have every likelihood of doing significantly better on this test than those of non-government schools who mostly belong to the lower socio-economic stratum.

To test if there is an equal variability among the students of government schools and those of non-government schools, the SE of the difference between the SD's was calculated by using formula 13 (vide page 75).The obtained difference between the SD's was .60 and the CR was found to be .63. This CR of .63 was far short of its value (1.96) at .05 level for 268 df. The obtained

difference was , therefore, not significant ($P > .05$). So, the rejection of the null hypothesis stands discredited on the basis of this test. Hence it seems reasonable to think that there is equal variability among the students of government schools and those of non-government schools with regard to their performance on this test.

CHAPTER FIVE

CORRELATIONAL STUDIES

This chapter presents an attempt to investigate whether our subjects will do better on NLTMA than on some standard tests of mental ability , and whether there exists any correlation between NLTMA and those tests.

ADMINISTRATION I

INTRODUCTION

The present study was carried out to investigate whether our subjects would do significatly better on NLTMA than on Raven's Standard Progressive Matrices Test (SPMT),and whether NLTMA was correlated with SPMT.

Hypothesis.

(1) Our subjects will do significantly better on NLTMA than on SPMT and (2) the former will have little correlation with the latter .

METHOD

Subjects.

A sample of 67 students of class X was drawn randomly by means of a table of random numbers (Dixon et al., 1951, pp. 290-294) from the population of students of class X in 3 High schools of Rajshahi Town in Bangladesh. The sample consisted of 65.67% boys, 34.33% girls, 38.81% Science and 61.19% Humanities students.

Fig. 4 graphically represents the distribution of smoothed frequencies of scores on SPMT. Fig. 5 shows the graphical representation of the distribution of smoothed frequencies of scores on NLTMA.

TABLE XXVII

Showing frequency distribution of NLTMA scores.

Group X-values	f	Smoothed f
70 and above	0	1.67
65-69	5	6.00
60-64	13	14.00
55-59	24	15.33
50-54	9	12.67
45-49	5	6.00
40-44	4	3.33
35-39	1	3.33
30-34	5	2.33
25-29	1	2.00
24 and below	0	0.33
	N = 67	

TABLE XXVIII

Showing frequency distribution of SPMT scores.

Group X-values	f	Smoothed f
50 and above	0	1.00
45-49	3	2.33
40-44	4	4.00
35-39	5	5.67
30-34	8	10.33
25-29	18	12.00
20-25	10	12.00
15-19	8	8.00
10-14	6	6.33
5 - 9	5	3.67
4 and below	0	1.67

N = 67

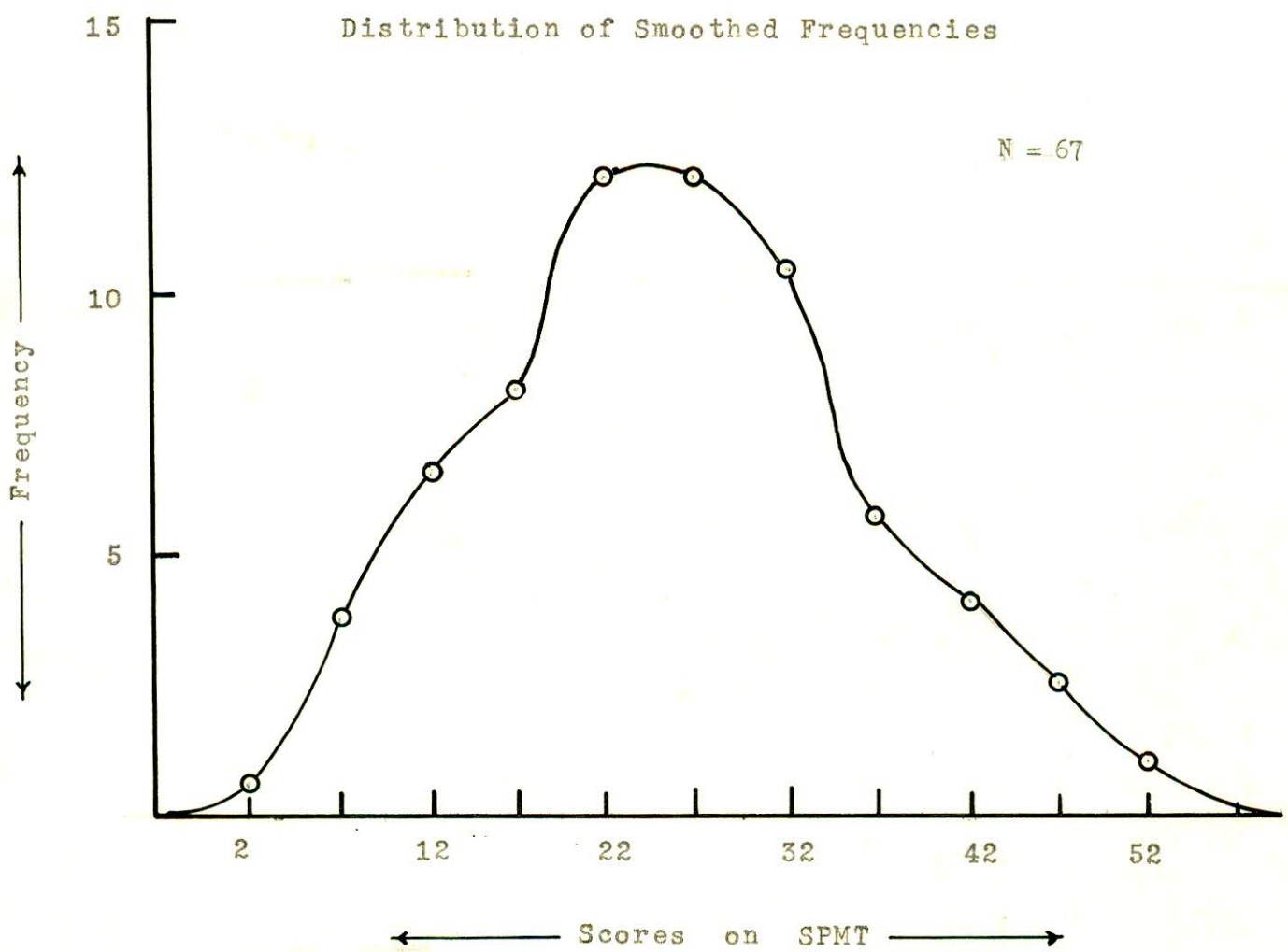


Fig.4

Frequency distribution of SPMT scores.

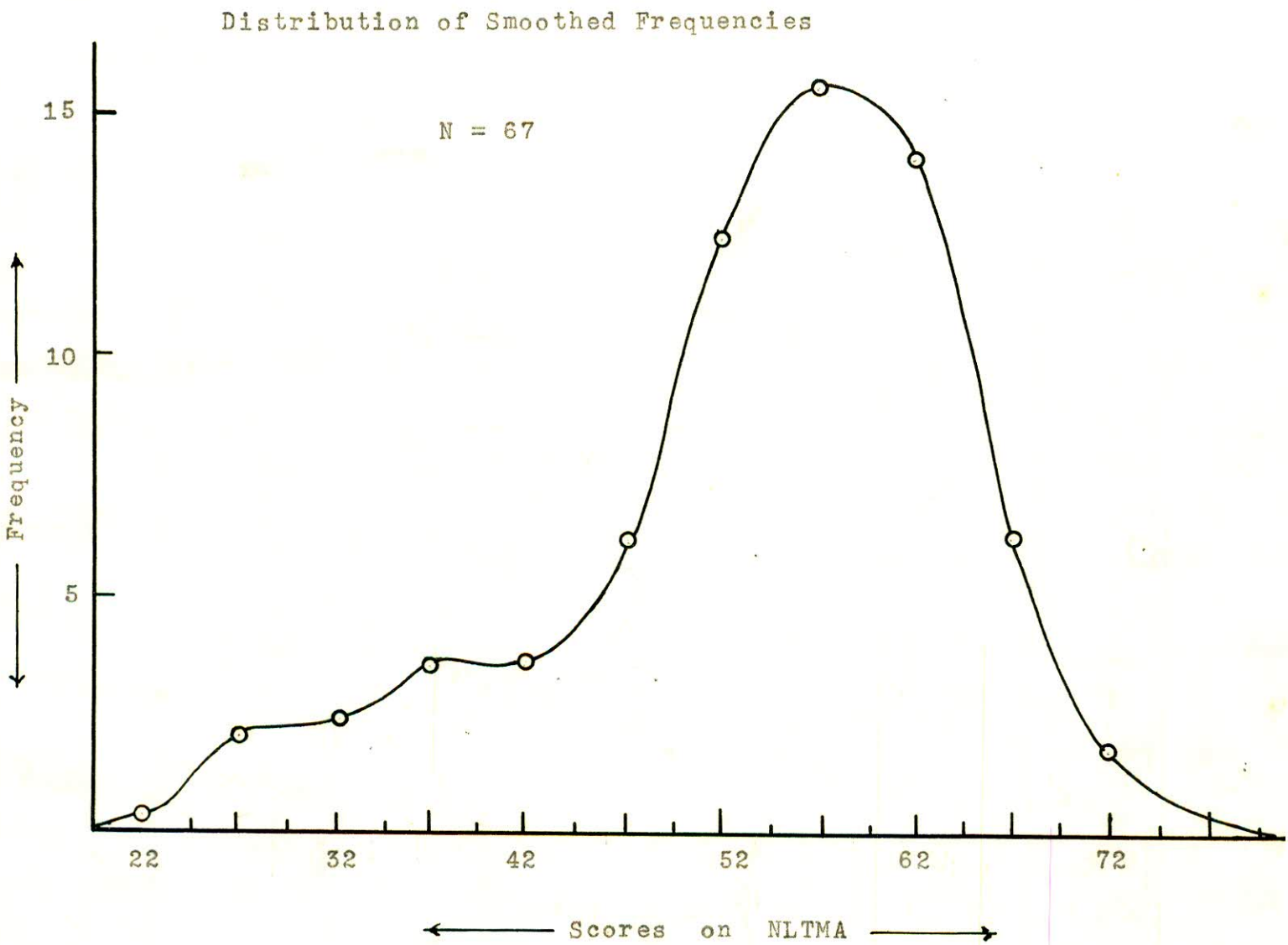


Fig.5

Frequency distribution of NLTMA scores.

TABLE XXIX

Showing comparison between mean performances on NLTMA and SPMT.

	Mean	SD	Variance	N	t-ratio
NLTMA	53.80	9.60	92.16		
				67	32.33
SPMT	25.35	10.25	105.06		

The PM correlation between NLTMA and SPMT scores was found to be 0.74, a significant ($P < .01$) positive correlation.

DISCUSSION

While comparing the mean performances of individuals on NLTMA and SPMT , it was found that t equaled 32.33 for 66 df. In consultation with the relevant table, we find that t reads approximately 2.00 and 2.65 at .05 and .01 levels respectively for 66 df. The obtained t of 32.33 far exceeds those values and is, therefore, significant ($P < .01$). It, therefore, seems reasonable to think that our subjects have every likelihood of doing significantly better on NLTMA than on SPMT.

The PM correlation coefficient between NLTMA and SPMT scores was found to be moderately high (0.74). This correlation coefficient was found to be

significant beyond .01 level for 65 df. ($P < .01$).

Because of the significant correlation between SPMT and NLTMA, a question may naturally be raised : Can they be treated as parallel tests ? But in view of the reasons advanced below, the two tests can-not be regarded as parallel. First, the correlation between SPMT and NLTMA is not as high as .90 or more. Most of the authors of standard intelligence tests report reliability coefficients of at least .90 between parallel forms of their tests (Garrett, 1960, p.351). The obtained correlation coefficient of .74 between SPMT and NLTMA falls far short of that value. Secondly, Ferguson speaks of parallel tests as having similar test contents, types of items,instructions for administering and the like for different forms. He further says that the parallel forms should have approximately equal means, equal standard deviations and equal intercorrelations (Ferguson, 1971,p.366). Gulliksen,too, says that the parallel tests have equal means, equal variances and equal intercorrelations with one another (Guilford, 1954,p.374; Gulliksen, 1950,pp.13-14). Anastasi maintains that parallel-form tests should contain the same number of items, and such items should be expressed in the same form and should cover the same type of content. She also says that the range and level of difficulty of the items should likewise be equal and that instructions,time limits,illustrative examples, format and all other aspects of the tests need to be checked for comparability (Anastasi, 1963, p. 119).

Judged by any criterion of parallel tests referred to above,the SPMT and NLTMA can-not be called parallel tests.

ADMINISTRATION II

INTRODUCTION

This study was undertaken to see if our subjects would do significantly better on NLTMA than on Terman's Non-Language Multi-Mental Test (NLMMT) Form A, and if NLTMA was correlated with NLMMT Form A.

Hypothesis.

(1) Our subjects will do significantly better on NLTMA than on NLMMT Form A and (2) the former will have little correlation with the latter.

METHOD

Subjects.

The same sample of 67 students of class X as used in administration I (vide page 79) was taken as subjects.

Tests used.

Non-Language Test of Mental Ability (NLTMA) and Terman's Non-Language Multi-Mental Test Form A (NLMMT).

Procedure.

After an interval of two days, the NLMMT Form A was administered to the same group of subjects on which NLTMA was previously applied. Subjects were seated one behind another in a room suited for the purpose. Care was taken to see that they felt at home and enjoyed the testing situation. They were praised and encouraged to have them motivated. They were asked to take all the time they needed and not

to make haste. They were also asked not to omit any item. When all the subjects finished answering the items in the manner suggested, the answer sheets were taken back. The scoring was performed as usual (Terman et al., 1942, p. 4). The significance of the difference between the mean performances on NLTMA and NLMMT Form A was statistically tested (t-test). The SE_D was estimated by using formula 14. The PM correlation between NLMMT and NLTMA scores was calculated by virtue of formula 7 (vide page 31).

RESULTS

The raw scores obtained on NLMMT Form A were translated into a frequency distribution(NLTMA frequency distribution has been shown in Table XXVII) as given in Table XXX. Fig. 6 graphically represents the distribution of smoothed frequencies of scores on NLMMT Form A.

TABLE XXX

Showing frequency distribution of NLMMT Form A scores.

Group X-values	f	Smoothed f
42 and above	0	.34
38-41	1	.67
34-37	1	2.33
30-33	5	9.67
26-29	23	15.33
22-25	18	19.00
18-21	16	11.67
14-17	1	6.33
10-13	2	1.00
9 and below	0	1.33
	N = 67	

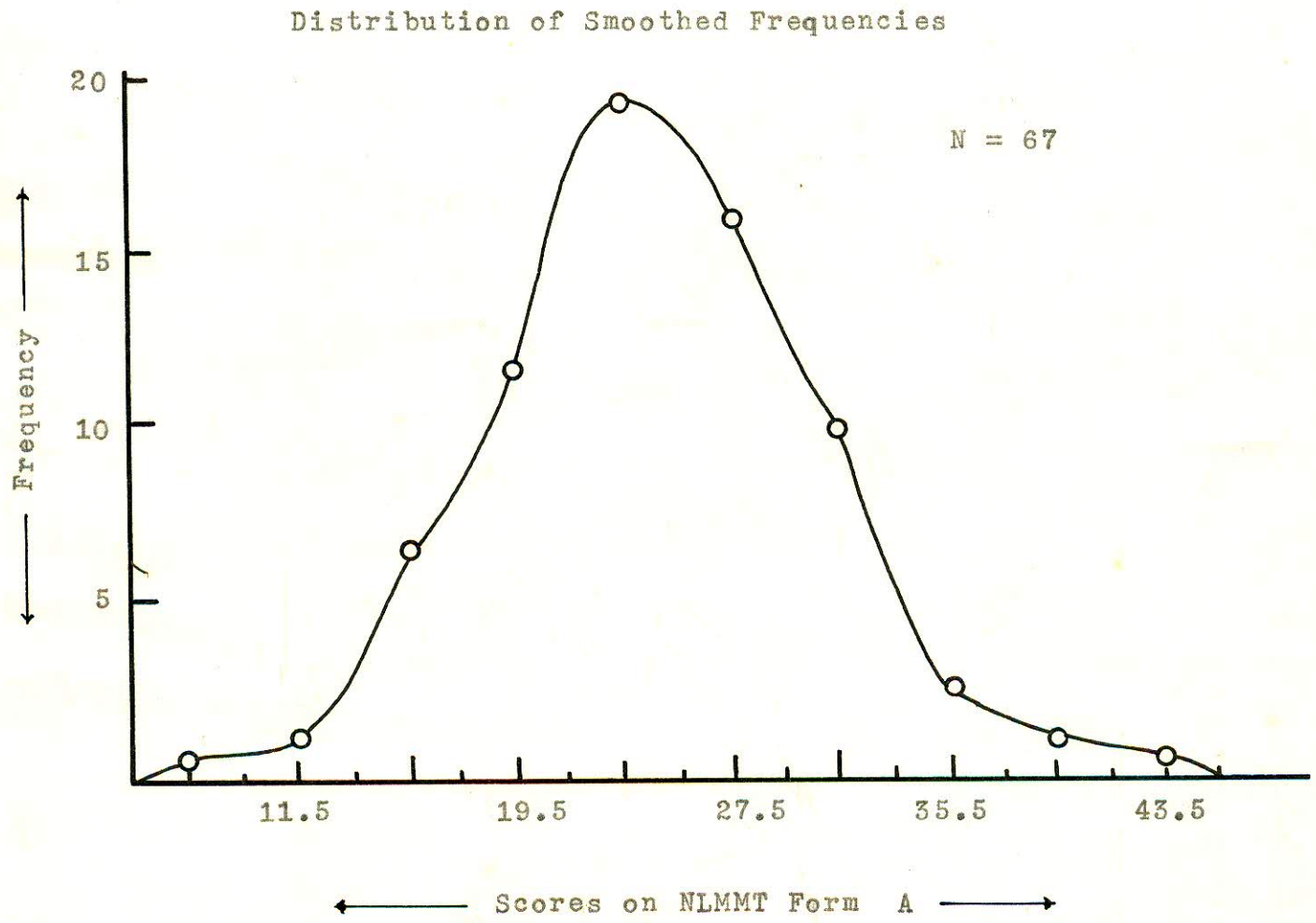


Fig. 6

Frequency distribution of NLMMT Form A scores.

TABLE XXXI

Showing comparison between mean performances on NLTMA and NLMMT Form A.

	Mean	SD	Variance	N	t-ratio
NLTMA	53.80	9.60	92.16		
				67	30.26
NLMMT Form A	27.47	4.88	23.81		

The PM correlation between NLTMA and NLMMT Form A scores was estimated to be 0.64, a significant ($P < .01$) positive correlation.

DISCUSSION

While comparing the mean performances of our subjects on NLTMA and NLMMT (Form A), it was found that t equaled 30.26. Referring to the table of t distribution, we find that t reads approximately 2.00 and 2.65 at 0.05 and 0.01 levels respectively for 66 df. The obtained t of 30.26 far exceeds those values and is, therefore, significant ($P < .01$). It, therefore, appears reasonable to think that there is every likelihood on the part of our subjects to do significantly better on NLTMA than on NLMMT.

The PM correlation between NLTMA and NLMNT Form A scores was found to be moderately high (0.64). This correlation coefficient was found to be significant beyond .01 level for 65 df ($P < .01$).

Because of the significant correlation between NLMNT and NLTMA, a question may naturally be raised : Can they be treated as parallel tests ? But in consideration of the reasons put forward below, these two tests can-not be taken as parallel. First, the obtained correlation coefficient is not as high as .90 or more . Most of the authors of standard intelligence tests report reliability coefficients of at least .90 between the parallel forms of their tests (Garrett, 1960,p.351).The obtained correlation between NLTMA and NLMNT falls far short of that value. Secondly, Ferguson speaks of parallel tests as having similar test contents, types of items,instructions for administering and the like for different forms. He also maintains that the parallel forms should have approximately equal means,equal standard deviations and equal intercorrelations (Ferguson,1971,p.366).Gulliksen says that the parallel tests have equal means,equal variances and equal intercorrelations with one another(Gulliksen ,1950,pp.13-14).Anastasi comments that the parallel-form tests should contain the same number of items,and such items should be expressed in the same form and should cover the same type of content. She further maintains that the range and difficulty

level of the items should likewise be equal and that instructions, time limits, illustrative examples, format, and all other aspects of the tests need to be checked for comparability (Anastasi, 1963, p.119).

Judged by any criterion of parallel tests mentioned above, NLTMA and *NLMMT Form A cannot be considered to be parallel.

---

* A good number of Raven's SPMT and Terman's NLMMT were available for purposes of comparison in the present study. Other non-language tests (e.g.,Cattell Culture-Fair Test, Penrose Pattern Perception Test, etc.) which also measure the spatial and perceptual abilities (vide pp.8-20 of this thesis) could not be used for the purpose as one or two copies of each were available and these again could not be duplicated for lack of technical facilities.

CHAPTER SIX

GENERAL DISCUSSION AND CONCLUSION.

It is perhaps reasonable to think now that the test which has been developed promises to serve the purpose for which it is intended. Some discussions of results have already been presented in different contexts. Nevertheless, some peculiar results need clarification.

The normality test shows that the frequency distribution of NLTMA scores does not differ significantly from the normal probability distribution (vide Table X) . This conformity in the distribution may imply that the intellectual ability assessed by the NLTMA in the population concerned is normally distributed. In other words, the population concerned may be normally distributed with respect to the intellectual ability measured by this test. The conformity may also mean that the sample was unbiased and that it represented the population from which the sample was drawn.

The reliability of the test was determined by the split-half (odd-even) method and was found to be .81. An odd-even reliability coefficient is usually lower than what is obtained by K-R Formula 20, since the latter takes into account all sorts of subdivisions of the test samples (Guilford, 1965, p. 460). The obtained coefficient alpha of .96 is evidently higher than the split-half (odd-even) reliability coefficient. This may mean that the present test consists of homogeneous items. Anastasi comments that 'unless the test items are highly homogenous, the K-R coefficient

(coefficient alpha) will be lower than the split-half reliability' (Anastasi, 1963, p. 123). The reliability coefficient determined by the test-retest method on the lapse of a period of 2 (two) months was found to be .68, which is much lower than those obtained by other methods. The parallel-form (alternate form) reliability technique is not applicable here since the NLTMA is a power test.(Guilford, 1965, pp. 450-452). The low test-retest reliability coefficient may mean that the individuals in the population have changed in different directions or in the same direction at different rates. These changes may be referred to as function fluctuation of individuals (Guilford,1965,p.451). In other words, the individual differences in the attributes assessed by NLTMA may depend upon time factor. The statistically insignificant ($P > .10$) difference between the initial test scores and the retest scores may indicate that the abilities (spatial and perceptual) as tapped by the test do not significantly change with passage of time. Longitudinal studies may, however, provide an answer as to whether these abilities will change positively or negatively with advance in time. It may be noted that Owens has demonstrated that the verbal abilities assessed by Army Alpha Form 6 significantly increase with increase in age, the time interval between the initial testing and retesting being 31 years (Owens, 1953, pp. 40-52).

The predictive validity of NLTMA was determined by the PM correlation of the test scores with the marks obtained by the subjects in the school Annual Examination. The correlation was found to be .45 - a significant ($P < .01$) validity coefficient , no doubt. This may imply that the test may be useful as a psychological instrument for predicting success of students in schools. The factorial validity of the test was determined by the complete centroid method of Professor L.L.Thurstone. The perceptual speed factor (factor a_1) was found to have +ve loadings on all the subtests. Hence this factor may be called a general factor. The second factor(factor a_2) may be called spatial visualization factor, which has maximum loadings on Test C (Series Completion test) and Test D (Pattern Synthesis test) and minimum loadings on Test A (Figure matching test) and Test B (Sign Patterning test). The significance of the loadings of factors a_1 and a_2 has been statistically tested. Results show that both the factors may be regarded as significant, excepting the loading of factor a_2 on Test A. Test A may, therefore, be looked upon as saturated with the factor a_1 . This is in agreement with the findings of other investigators(Guilford,1968,p.243). The maximum loadings of .604 and .670 with factor a_2 on Test C and Test D respectively may mean that an individual has to preceive relations as well as handle objects imaginally in space to find solutions to the problems in those tests. The loadings of .416 and .420 with factor a_1 on Test C and Test D respectively may indicate that the individual has to perceive relations with

rapidity and accuracy to solve the problems in these tests. Again, Test A and Test B are loaded maximally (.642 and .639 respectively) with factor a_1 (perceptual factor), and minimally (-.032 and .144 respectively) with factor a_2 (spatial visualization factor) (vide page 60). This may imply that an individual requires mostly to perceive relations with rapidity and accuracy, and very little to handle objects imaginally in space, to solve the problems in those tests. Factor a_2 may be regarded as bipolar in nature, since it has both +ve and -ve loadings on the subtests. It may be noted from Table XIX (vide page 61) that factor a_1 has taken 57% and factor a_2, 43% of the total factor variances.

Table XXIV (vide page 74) shows that boys have done significantly ($P < .01$) better than girls on the test. In other words, boys are superior to girls in respect of their performance on the abilities measured by the present test. As the test assesses the factors of spatial visualization and perceptual speed, it may be concluded that boys are better than girls in respect of these abilities. These results support the findings of other workers(Guilford, 1967,pp.404-406). The superior performance of the boys may be attributed to familiarity factor. Boys may be more familiar with the test items than girls. Boys enjoy more freedom of getting about here and there and as such, may acquire more familiarity with space than girls. Table XXIV apparently shows that there is greater variability among boys than among girls. But the statistical test (t-test) does not support this ($P > .05$).

It may be further noted from Table XXV (vide page 75) that Science students have done significantly ($P < .01$) better than Humanities students on NLTMA. This may be due to the fact that Science students are more familiar with geometrical figures than Humanities students. Another reason for it may be that better students usually go to science classes. It is also seen that there is lesser variability among science students than among Humanities students. But the statistical test(t-test) does not corroborate this($P > .05$).

Table XXVI (vide page 77) apparently shows that there is less variability among students of government schools than among those of non-government schools. But the statistical test (t-test)does not support this ($P > .05$).Again, the former have been found to differ significantly ($P < .01$) from the latter in respect of their performance on NLTMA. A reason for it may be that better students usually go to such schools. Another reason may be that the students of government schools who mostly belong to the upper socio-economic stratum of the society are environmentally superior to those of non-government ones who mostly belong to the lower socio-economic stratum of the society. In other words, the significant difference in the mean performance on the test may be considered to be due to environmental facilities and better calibre of students.Incidentally, it may be mentioned that government schools are equipped with better teaching facilities. However, corroboration of this by further studies is necessary. The present study does not claim to provide

any answer to this problem.

An attempt was made to investigate if NLTMA correlated with Raven's Standard Progressive Matrices and Terman's Non-Language Multi-Mental Test. The former was found to correlate significantly ($P < .01$) with SPMT and NLMMT Form A, the correlations being .74 and .64 respectively. This may mean that they measure some common factors. Factor analysis has revealed that the present test may particularly measure spatial visualization factor and perceptual speed factor. Raven's Standard Progressive Matrices Test (SPMT) measures a general factor and a spatial factor (Raven, 1951, p.1), and Terman's Non-Language Multi-Mental Test (NLMMT) Form A mainly evaluates the perceptual speed factor (Terman, et al., 1942, p.1). This may explain the correlation between them. Again , a comparison of the mean performances on SPMT, NLMMT and NLTMA has shown that our subjects have done significantly ($P < .01$) better on NLTMA than on both SPMT and NLMMT Form A. This may mean that this test is more suited to our culture. It may be further noted that the significant ($P < .01$) correlation of the present test with SPMT and NLMMT may be taken as an additional support for the validity of the former.

The facts presented in this dissertation with regard to the NLTMA are principally based on statistical observations. With this point in mind, the following conclusions may be drawn :

1. The NLTMA may be regarded as a nonverbal test having to do with the assessment of spatial and perceptual abilities.
2. The present test may be regarded as free from ambiguity in respect of interpretation of scores because of the high coefficient alpha it possesses.
3. This test has been developed through administration on subjects in our culture, and they have been found to do significantly better on it than on certain nonverbal tests (SPMT and NLMMT). On the other hand, different psychologists have constructed nonverbal tests in countries abroad but have failed to show that these are culture-free. Hence the present test may be used in our country as a substitute for those tests.
4. The present test may be of use in the prediction of success in schools .
5. Boys and girls have every likelihood of performing differentially on this test.
6. The abilities assessed by this test may change very little with passage of time. Longitudinal studies on this test may, however, be carried out to support this point.
7. That the test assesses the intellectual ability of an individual gets an additional support from the significant ($P < .01$) correlations that it bears with Raven's SPMT and Terman's NLMMT.
8. The present test may be of use in assessing the intellectual ability of individuals whose performance is poor on verbal tests,or who are illiterate or physically handicapped.Further studies on the test on samples of such individuals should, however, be undertaken to support this point.

BIBLIOGRAPHY

Anastasi, Anne. Psychological Testing. New York: Macmillan, 1963.

Ansbacher, H.L. The Goodenough Draw-a-Man Test and Primary Mental Abilities, J.Consult. Psychol., 1952, 16, 176-180.

Anstey, E. Psychological Tests. London : Nelson, 1966.

Arthur, G. The Arthur Adaptation of the Leiter International Performance scale. Washington : Psychol. Ser. Press, 1952.

Bennett, G.K., Bennett, Marjorie G., Wallace, W.L., and Wesman, A.G. College Qualification Tests. New York: Psychol. Corp., 1961.

Bradway, K.P., Thompson, C.W., and Cravens, R.B. Preschool IQ's after 25 years, J.Educ. Psychol., 1958, 49, 278-281.

Brown, A.W., Rohrer, P.T., and Stein, S.P. The Chicago Non-verbal Examination. New York : Psychol. Corp., 1936-1947.

Buros, O.K. (Ed.) The Fourth Mental Measurement Year Book. New Jersey : Gryphon, 1953.

Buros, O.K. (Ed.) The Fifth Mental Measurement Year Book. New Jersey : Gryphon, 1959.

Burt, C. The Evidence for the concept of Intelligence, Brit.J.Educ. Psychol., 1955, 25, 1, 158-177.

Burt, C. Factor Analysis and its Neurological Basis, Brit.J.Statist. Psychol., 1961, 14, 1, 54-71.

Butcher, H.J. Human Intelligence. London : Methuen, 1968.

Cattell, R.B. A Culture-Free Test of Intelligence, I.J.Educ. Psychol., 1940, 3, 161-179.

Cattell,R.B. and Cattell, A.K.S.Handbook for the Individual or Group Culture-Fair Intelligence Test. Champaign,Illinois : Inst.Pers.Abil. Test., 1959.

Cattell, R.B. The Theory of Fluid and Crystallized General Intelligence checked at 5-6 year level,Brit.J.Educ. Psychol..1967,37,2,209-223.

Cohen,J.The Factorial structure of the WAIS between early adulthood and old age, J.Consult.Psychol..1957,21,283-290.

Cornell,E.L.and Coxe,W.W. A Performance Ability Scale. Tarry town-on Hudson, N.Y. : World Book, 1934.

Cronbach, L.J.Coefficient alpha and the internal structure of tests, Psychometrika. 1951,16, 299-334.

Cronbach,L.J.Essentials of Psychological Testing. New York: Harper,1960.

Davis, A.and Eells, K.Davis-Eells Games : Davis-Eells Test of General Intelligence or Problem-solving Ability.Tarry town-on Hudson,N.Y.: World Book, 1953.

DeCarre, T.G. Intelligence and Affectivity in Early Childhood. New York: International University Press, 1965.

Dixon,W.J.and Massey, F.J.Jr.Introduction to Statistical Analysis. New York : McGraw-Hill, 1951.

Ferguson,G.A.Statistical Analysis in Psychology and Education.New York : McGraw-Hill, 1971.

Fisher,R.A.Statistical Methods for Research Workers. London: Oliver and Boyd, 1967.

Flanagan,J.C.General Consideration in the selection of test items and a short method of estimating the product-moment coefficient from the data at the tails of a distribution,J.Educ.Psychol.,1939,30,674-680.

Forbes,A.R.An item analysis of the Advanced Matrices,Brit.J.Educ.Psychol.,1964, 34, 3, 223-236.

Freeman, F.S.Theory and Practice of Psychological Testing.(Revised Edition). New York : Holt,Rinehart and Winston, 1960.

Garrett, H.E.Statistics in Psychology and Education. London : Longmans,1960.

Ghiselli,E.E.Theory of Psychological Measurement. New York:McGraw-Hill,1964.

Gibson,W.A.An Asymmetric Approach to Multiple Factor Analysis,Brit.J.Statist. Psychol., 1961,14, 2, 98-107.

Goodenough,F.L. Measurement of Intelligence by drawings. Tarry town-on-Hudson, N.Y.: World Book, 1926.

Guilford,J.P.Psychometric Methods. New York:McGraw-Hill,1954.

Guilford,J.P.Fundamental Statistics in Psychology and Education. New York : McGraw-Hill, 1965.

Guilford,J.P.The Nature of Human Intelligence. New York:McGraw-Hill, 1967.

Guilford,J.P.The Structure of Intelligence. In Dean K.Whitla (Ed.),Handbook of Measurement and Assessment in Behavioural Sciences.London: Addison-Wesley, 1968,pp. 215-260.

Gulliksen,H.Theory of Mental Tests. New York: Wiley, 1950.

Guttman,L. The structure of interactions among intelligence tests. Princeton, N.J. : Educational Testing Service, 1965.

Harris,D.B. and Goodenough,F.L.Harris-Goodenough Measures of Intellectual Maturity. Tarry town-on-Hudson,N.Y.:Harcourt-Brace,1963.

Hays,W.L.Statistics for Psychologists. New York: Holt,Rinehart and Winston,1963.

Hebb,D.O.A Text Book of Psychology, Tokyo:Tuttle,1967.

Holm,F.E.Elementary Matrix Algebra. New York:Macmillan-Collier,1964.

Horst,P.Matrix Algebra for Social Scientists. New York:Holt,Rinehart and Winston, 1963.

Hunt,J.McV.Intelligence and Experience. New York:Ronald, 1961.

Jarrett,R.F. A Note on Rotation and Psychological Space,Brit.J.Statist.Psychol., 1961, 14, 2, 109-115.

Jones,L.V.A Factor Analysis of the Stanford-Binet Scale at 4-yr.age level, Psychometrika. 1949, 14, 299-331.

Kelley,T.L.The Selection of Upper and Lower Groups for the validation of items, J.Educ. Psychol., 1939, 30, 17-24.

Knight,R.Intelligence and Intelligence Tests. London: Methuen,1953.

Lawley,D.N.and Maxwell,A.E.Factor Analysis As a Statistical Method. London: Butterworths, 1963.

Littell,W.M.The Weschler Intelligence scale for Children-A Review of a decade of research, Psychol. Bull., 1960, 57,132-156.

Mac Murray,D.A.A comparison of gifted children and of dull-normal children measured by the Pintner-Paterson Scale,as against the Stanford-Binet Scale, J.Psychol., 1937, 4, 273-280.

McCall,W.A.Measurement, New York: Macmillan, 1939.

McCarthy,D.A study of the reliability of the Goodenough Test of Intelligence, J.Psychol.,1944, 18,201-216.

McNemar,Q.Psychological Statistics. New York: Wiley, 1949.

Morgan,C.T.and Stellar,E.Physiological Psychology. New York:McGraw-Hill,1950.

Morris,C.M.A critical Analysis of certain Performance Tests,J.Genet. Psychol., 1939,54,85-105.

Mosier,C.I.A short cut in the estimation of split-half Coefficients,Educ.Psychol. Measmt.,1941, 1, 407-408.

Mursell,J.L.Psychological Testing. New York: Longmans, 1947.

Nunnally, Jum C. Tests and Measurements.Assessment and Prediction. New York: McGraw-Hill, 1959.

Nunnally, Jum C.Psychometric Theory. New York: McGraw-Hill, 1967.

Otis,A.S.Otis Self-Administering Test of Mental Ability. New York: Harcourt-Brace, 1928.

Owens,W.A.Age and mental abilities:a longitudinal study,Gen.Psychol.Monogr., 1953, 48,3-54.

Patterson,R.M.The significance of practice effect upon readministration of the Grace Arthur performance scale to High-grade mentally deficient children, Amer.J.Ment.Defic.,1946,50,393-401.

Pearson,E.S.and Hartley,H.O.(Ed.) Biometrika Tables for Statisticians.Vol.1. Cambridge:Cambridge University Press,1966.

Piaget,J.The Psychology of Intelligence. London: Routledge and Kegan Paul,1951.

Pintner,R.and Paterson,D.A Scale of Performance Tests. New York : Appleton, 1917.

Raven,J.C.Progressive Matrices. Dumfries : Crichton Royal,1951.

Richardson,C.A. An Introduction to Mental Measurement and its Application. London: Longmans, 1955.

Richardson,M.W.and Kuder,G.F.The Calculation of Test reliability Coefficient based on the method of rational equivalence,J.Educ.Psychol.,1939,30,681-687.

Rimoldi,H.J.A. A Note on Raven's Progressive Matrices Test,Educ.Psychol. Measmt., 1948, 8, 347-352.

Russell,W.R.Brain.Memory.Learning. Oxford: Clarendon,1959.

Spearman,C.The Abilities of Man.New York: Macmillan,1927.

Stoddard,G.D.The Meaning of Intelligence. New York: Macmillan,1956.

Sultan,E.E. A Factorial study in the domain of creative thinking,Brit.J. Educ. Psychol. 1962,32,78-82.

Tate,M.W.Statistics in Education. New York: Macmillan, 1955.

Terman,L.M.and McNemar,Q. Terman-McNemar Test of Mental Ability. New York: World Book, 1942.

Terman,E.L.,McCall,W.A.and Lorge,I. Non-Language Multi-Mental Test.Columbia University Teachers' College,1942.

Terman,L.M.Concept Mastery Test. New York: Psychol.Corp.,1956.

Thomson,G. The Factorial Analysis of Human Ability. London: University of London Press,1951.

Rajshahi University Library
Document Section
Document No. D-1044
Date. 2/3/88

Thurstone,L.L. The Nature of Intelligence. New York: Harcourt-Brace,1926.

Thurstone,L.L.and Thurstone,T.G. Tests of Primary Mental Abilities. Chicago: Science Research Associates, 1938-1950.

Thurstone,L.L.and Thurstone,T.G. Psychological Examination for College Freshmen. New York: Educational Testing Service, 1954.

Thurstone,L.L.Multiple Factor Analysis. Chicago: Chicago University Press, 1947.

Tucker,H.G. An Introduction to Probability and Mathematical Statistics. New York: Academic Press, 1965.

Vernon,P.E. The Structure of Human Abilities. London: Methuen, 1950.

Weatherburn,C.E. A First Course in Mathematical Statistics. London: Cambridge University Press, 1962.

Weschler,D. The Measurement of Adult Intelligence. Baltimore: Williams and Wilkins, 1944.

Weschler,D. Weschler Intelligence Scale for children. New York : Psychol.,Corp., 1949.

Weschler,D. Weschler Adult Intelligence Scale. New York : Psychol. Corp.,1955.

Woodworth,R.S. Contemporary Schools of Psychology. London: Methuen, 1960.

Wyatt,H.G. The Psychology of Intelligence and Will. London: Kegan-Paul, 1930.

APPENDIX

Instructions for Test A

Instructions for Test B

Instructions for Test C

Instructions for Test D

Instructions in Bengali for Test A

Instructions in Bengali for Test B

Instructions in Bengali for Test C

Instructions in Bengali for Test D

Scoring key to Test A

Scoring key to Test B

Scoring key to Test C

Scoring key to Test D

Answer sheet for Test A

Answer sheet for Test B

Answer sheet for Test C

Answer sheet for Test D

Test A Figure matching test

INSTRUCTIONS

Please go through the following instructions carefully and do accordingly.

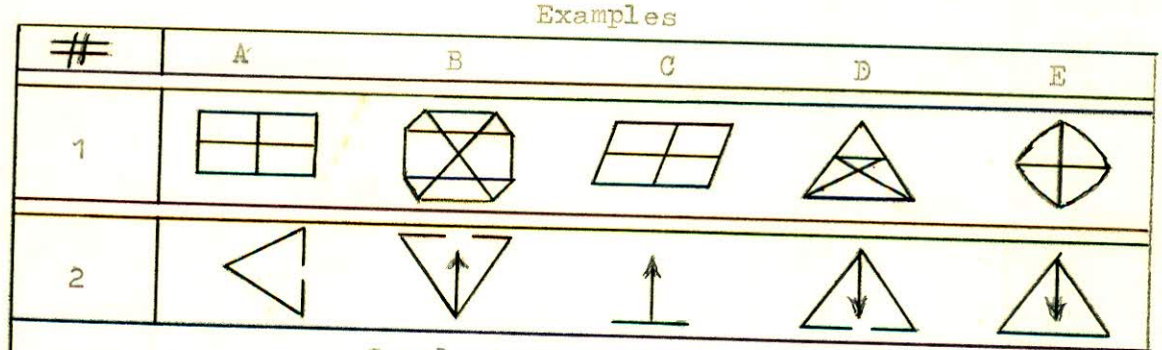
In this test you will find that each row has five (A,B,C,D,E) figures,the two of which are similar in certain respects. In every row you are to look carefully at each figure and to find the two which are similar in some respects. You are to indicate your choices by putting cross-marks (X) in the squares below the two letters (A,B,C,D,E) which top the correct figures in the answer sheet supplied.

There are two practice items below. In the first example, the figures below the letters A and C are similar ; because, both of them are quadrilaterals (rectangle and parallelogram) and two straight lines have joined the midpoints of their lengths and breadths. In the second example, the figures below the letters B and D are similar ; because, though one is an inverted figure of the other, their vertical angles have been bisected by arrowheads. You find below in the sample answer sheet that the squares below the letters AC and BD respectively for the two practice items (1&2) have been cross-marked.

Please examine the figures from 1 to 20 on the next page (vide test booklet) and find the correct answers. Please indicate the correct choices in the answer sheet by putting two cross-marks in the squares below the letters containing the figures which are similar in some respects.

Please do not omit any item and take all the time you need. Please do not make any mark in this booklet. Please do not turn over the page until told to do so.

Examples



Sample Answer sheet

#	A	B	C	D	E
1	☒	☐	☒	☐	☐
2	☐	☒	☐	☒	☐

Test B Sign patterning test

INSTRUCTIONS

Please go through the following instructions carefully and do accordingly.

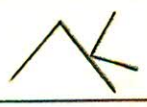
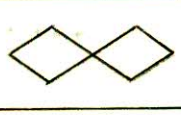
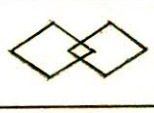
In this test you will find that each row has four (A,B,C,D) figures in addition to one at the extreme left. The figure (a combination of two or more figures) at the extreme left has some sign or signs put into it in a certain way. In each row you are to look carefully at every figure and to find one from among the four (A,B,C,D) in which the same sign or signs may be placed in the same way. You are to indicate your choice by putting a cross-mark (X) in the square below one of the letters (A,B,C,D) topping the correct answer figure in the answer sheet supplied.

There are two practice items below. In the first example, two angles have met together so as to make a space. A plus sign has been placed in the middle of the space in front of the angles. In the second example, two quadrilaterals have met together anglewise so as to make a space. A plus sign has been placed in the middle of the space in front of the angles. Hence you find below in the sample answer sheet that the squares below the letters A and C respectively for the two practice items (1&2) have been cross-marked (X).

Please examine figures from 1 to 25 on the following page(vide test booklet) and find the correct figures. Please indicate your choice in the answer sheet by putting a cross-mark (X) on the square below one of the letters (A,B,C,D).

Please do not omit any item and take all the time you need. Please do not make any mark on this booklet. Please do not turn over the page until told do so.

Examples

#		A	B	C	D
1					
2					

Sample Answer sheet

#	A	B	C	D
1	☒	☐	☐	☐
2	☐	☐	☒	☐

Test C

Series completion test

INSTRUCTIONS

Please go through the following instructions carefully and do accordingly.

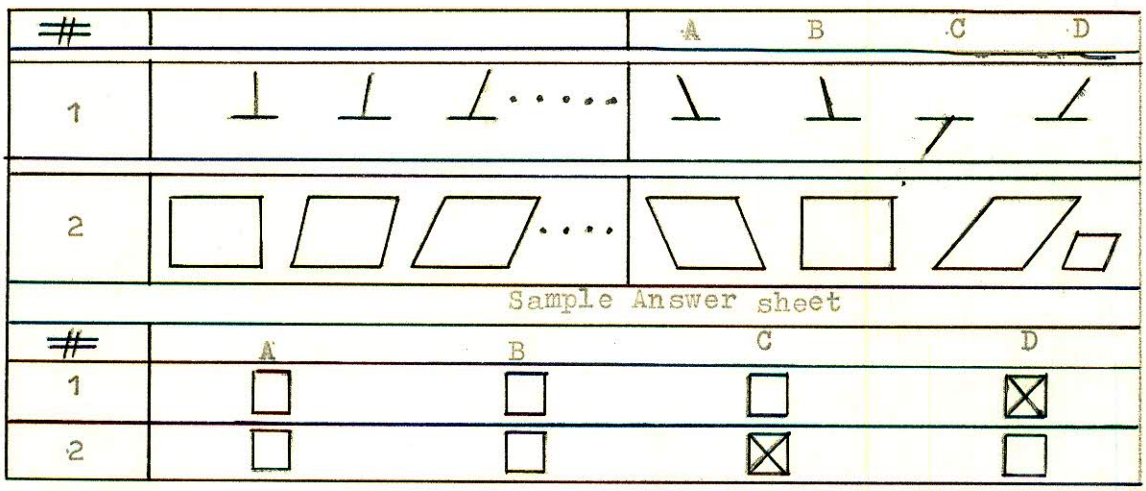
In this test you will find that each row has four (A,B,C,D) figures in addition to three at the left. Every row at the left has a gap indicated by dots. In each row on the right you are to look carefully at every figure and to find one which will complete the series at the left. You are to indicate your choice by putting a cross-mark(X) in the square below one of the letters (A,B,C,D) topping the correct figure in the answer sheet supplied.

There are two practice items below. In the first example, you find in the first of the lefthand figures that a straight line is perpendicular to the other. In the second figure, the perpendicular is inclined to the right and in the third, it is more inclined to the right. So, in the figure on the dotted space, the perpendicular will be much more inclined to the right. As such, the correct figure stands below the letter D. In the second example, you find that the first of the left-hand figures is a rectangle. In the second figure the perpendicular lines are inclined to the right and in the third, they are more inclined to the right. So, in the figure on the dotted space, they will be much more inclined to the right. As such, the correct figure stands below the letter 'C'. You will, therefore, find that the squares below the letters D and C respectively for the two practice items (1&2) are cross-marked in the sample answer sheet below.

Please examine the figures from 1 to 20 on the following page (vide test booklet) and find the correct answer. Please indicate your choice in the answer sheet by putting a cross-mark (X) in the square below one of the letters (A,B,C,D) containing the correct figure.

Please do not omit any item and take all the time you need. Please do not make any mark on this booklet. Please do not turn over the page until told to do so.

Examples



Sample Answer sheet

#	A	B	C	D
1	☐	☐	☐	☒
2	☐	☐	☒	☐

Test D Pattern synthesis test

INSTRUCTIONS

Please go through the following instructions carefully and do accordingly.

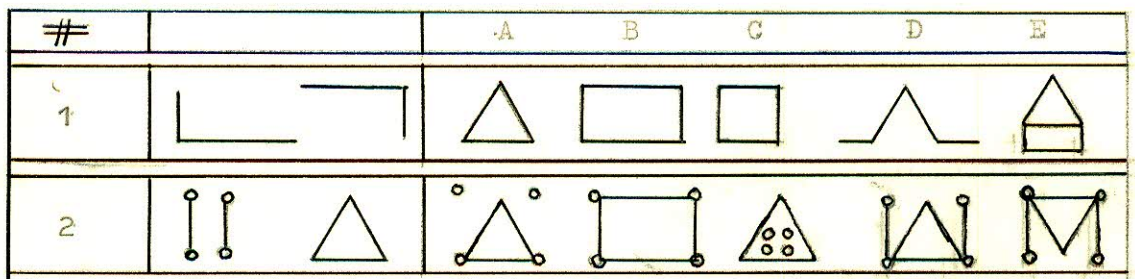
In this test you will find that each row has five (A,B,C,D,E) figures in addition to two or more at the left. In every row at the right you will find that the figures (A,B,C,D,E) are the combination of two or more figures at the left. You are to look carefully at each figure on the right and to find one which results from combining two or more figures at the left. You are to indicate your choice by putting a cross-mark (X) in the square below one of the letters (A,B,C,D,E) containing the correct figure in the answer sheet supplied.

There are two practice items below. In the first example, you find that the figures below the letters B and C may both be formed by the combination of two figures at the left. But if you carefully look at the figure below C, you will find that it is smaller. Hence, the correct figure stands below the letter B. In the second example, you find that the figures below the letters D and E may both be formed by the combination of three figures at the left. But if you carefully look at the figure below E, you will find that the figure below E is inverted. So, the correct figure stands below the letter D. Hence the squares below the letters B and D for the two practice items (1&2) respectively are cross-marked in the sample answer sheet below.

Please examine the figures from 1 to 20 on the next page (vide test booklet) and find the correct choices. Please indicate your choice in the answer sheet by putting a cross-mark (X) in the square below one of the letters containing the correct choice.

Please do not omit any item and take all the time you require. Please do not make any mark on this booklet. Please do not turn over the page until told to do so.

Examples



Sample Answer sheet

#	A	B	C	D	E
1	☐	☒	☐	☐	☐
2	☐	☐	☐	☒	☐

টেষ্ট ক

নির্দেশাবলী

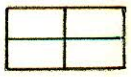
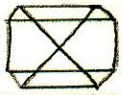
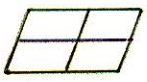
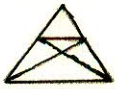
নীচে লেখা নির্দেশগুলো পড়ে দেখ । এ টেষ্টে দেখবে প্রতি সারিতে পাঁচটি (ক, খ, গ, ঘ, ঙ) করে ছবি আঁকা রয়েছে । এর ভিতরে দুটি ছবিতে কিছু মিল রয়েছে । প্রতি সারির প্রত্যেক ছবি মনোযোগ সহকারে দেখে যে দুটি ছবিতে মিল রয়েছে তাদের খুঁজে বের করো । পরে উত্তর - পত্রে তোমার বেছে নেওয়া ছবি দুটির ঠিক উপরে লেখা অক্ষর (ক, খ, গ, ঘ, ঙ) দুটির নীচের খোপে ক্রস - চিহ্ন দাও ।

নীচে দুটি উদাহরণ দেওয়া আছে । প্রথম উদাহরণে ক ও গ অক্ষর দুটির নীচের ছবি দুটিতে মিল রয়েছে । কারণ, ছবি দুটি চতুর্ভুজ (আয়তক্ষেত্র এবং সামান্তরিক) এবং উভয়ের দৈর্ঘ্যের ও প্রস্থের মধ্যবিন্দুদ্বয় দুটি সরলরেখা দিয়ে যোগ করা হয়েছে । দ্বিতীয় উদাহরণে খ ও গ অক্ষর দুটির নীচের ছবি দুটিতে মিল রয়েছে । কারণ, বিপরীতভাবে অবস্থিত হলেও উভয়ের শীর্ষকোণদ্বয় তীর-চিহ্নিত সরলরেখা দিয়ে দ্বিখণ্ডিত করা হয়েছে । তাই নীচের নমুনা উত্তর - পত্রে প্রথম উদাহরণে সঠিক উত্তর - সূচক অক্ষরদুটি ক ও গ এর এবং দ্বিতীয় উদাহরণে খ ও ঘ এর নীচের খোপ দুটিতে ক্রস-চিহ্ন (X) দেওয়া হয়েছে ।

পরবর্তী পৃষ্ঠায় ১ নং থেকে ২০ নং সমস্যা গুলো (অতীহা - পত্র দ্রষ্টব্য) মনোযোগ দিয়ে দেখ এবং ঠিক উত্তর বের করো । পরে উত্তর-পত্রে তোমার বেছে নেওয়া ছবি দুটির নীচের খোপে ক্রস-চিহ্ন (X) দাও ।

কোন সমস্যা বাদ দিও না । যতক্ষণ পার চেষ্টা করো । এ পুস্তিকায় কোন দাগদিও না । না বলা পর্য্যন্ত পাতা উল্টাইও না ।

উদাহরণ

#	ক	খ	গ	ঘ	ঙ
১					
২					

নমুনা উত্তর-পত্র

#	ক	খ	গ	ঘ	ঙ
১	☒	☐	☒	☐	☐
২	☐	☒	☐	☒	☐

টেস্ট ৪

নির্দেশাবলী

নীচে লেখা নির্দেশ খুঁটিয়ে পড়ে দেখ । এ টেস্টে দেখবে প্রতি সারিতে ডানদিকে চারটি (ক,খ,গ,ঘ) এবং বামদিকে একটি করে ছবি (যাহা একাধিক ছবির মিলনে গঠিত) আঁকা রয়েছে । এক বিশেষ নিয়ম অনুসারে বামদিকের ছবিতে এক বা একাধিক চিহ্ন দেওয়া হয়েছে ।

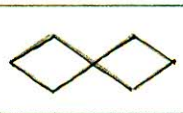
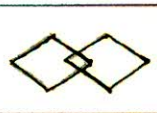
ডানদিকের প্রতি সারির ছবিগুলো বিশেষভাবে লক্ষ্য করে দেখ । তারপরে ডানদিকের ছবিগুলির ভিতর থেকে এমন একটি ছবি খুঁজে বের করো যাতে একই নিয়মে চিহ্ন বা চিহ্নগুলি দেওয়া যায় । পরে উত্তর-পত্রে তোমার বেছে নেওয়া ছবিটির উপরে লেখা অক্ষরটির (ক,খ,গ,ঘ) নীচের বোপে একটি ক্রস-চিহ্ন (X) দাও ।

দেখ , নীচে দুটি উদাহরণ দেওয়া আছে । প্রথম উদাহরণে বামদিকের ছবিতে দুটি কোণ পরস্পরের সাথে মিলেছে । কোণ দুটির সম্মুখে ছবিটির বামখানে একটি যোগ-চিহ্ন (+) দেওয়া আছে । দ্বিতীয় উদাহরণে বামদিকের ছবিতে দুটি চতুর্ভুজ কোণাকোণি ভাবে মিলায় শূণ্যস্থানের সৃষ্টি হয়েছে । উক্ত শূণ্যস্থানের মাঝ খানে কোণ দুটির সামনে একটি যোগ-চিহ্ন দেওয়া আছে । তাই নীচের নমুনা উত্তর-পত্রে প্রথম উদাহরণে নির্ভুল উত্তর-সূচক অক্ষরটি ক এর এবং দ্বিতীয় উদাহরণে গ এর নীচের বোপ দুটিতে ক্রস-চিহ্ন (X) দেওয়া হয়েছে ।

পরবর্তী পৃষ্ঠায় ১ নং থেকে ২৫ নং সমস্যা গুলো (অতীক্ষা-পত্র দ্রষ্টব্য) মনোযোগ দিয়ে লক্ষ্য করো এবং নির্ভুল উত্তর বের করো । পরে উত্তর-পত্রে তোমার বেছে নেওয়া ছবিটির উপরে লেখা অক্ষরটির (ক,খ,গ,ঘ) নীচের বোপে ক্রস-চিহ্ন (X) দাও ।

কোন সমস্যা বাদ দিও না । যতদূর পার চেষ্টা করো । এ পুস্তিকায় কোন দাগ দিও না । না বলা পর্য্যন্ত পাতা উল্টাইও না ।

উদাহরণ

#		ক	খ	গ	ঘ
১					
২					

নমুনা উত্তর-পত্র

#	ক	খ	গ	ঘ
১	☒	☐	☐	☐
২	☐	☐	☒	☐

টেস্ট ৭

নির্দেশাবলী

নীচে লেখা নির্দেশ গুলো পড়ে দেখ । এ টেস্টে দেখবে প্রতি সারিতে ডানদিকে চারটি (ক,খ,গ,ঘ) এবং বামদিকে তিনটি করে ছবি আঁকা রয়েছে । বামদিকের প্রত্যেক সারিতে কিছুটা শূণ্যস্থান (বিন্দু-চিহ্ন দেওয়া) রয়েছে । ডানদিকের ছবি চারটির ভিতর থেকে এমন একটি ছবি খুঁজে বের করো যাকে বামদিকের সারির শূণ্যস্থানে বসালে একটি সিরিজ বা ধারা সম্পূর্ণ হয় । নির্ভুল উত্তর-সূচক ছবিটি বেছে নিয়ে উত্তর-পত্রে উহার উপরিস্থিত অক্ষরটির (ক,খ,গ,ঘ) নীচের খোপে একটি ক্রস-চিহ্ন (X) দাও ।

নীচে দুটি উদাহরণ দেওয়া আছে । প্রথম উদাহরণে বামদিকের প্রথম ছবিতে একটি সরলরেখা অপরটির উপর নম্র । দ্বিতীয় ছবিতে নম্রটি ডানদিকে হেলে পড়েছে । তা'হলে শূণ্যস্থানের ছবিতে নম্রটি আরো বেশী হেলে পড়বে । তাই সঠিক উত্তর-সূচক অক্ষরটি ঘ । দ্বিতীয় উদাহরণে বামদিকের প্রথম ছবিটি আয়তক্ষেত্র । দ্বিতীয় ছবিতে আয়তক্ষেত্রের খাড়া বাহু দুটি ডানদিকে হেলে পড়েছে এবং তৃতীয় ছবিতে ডানদিকে আরো হেলে পড়েছে । তাই শূণ্যস্থানের ছবিতে বাহু দুটি ডানদিকে আরো বেশী হেলে পড়বে । সুতরাং সঠিক উত্তর-সূচক অক্ষরটি গ । তাই নীচের নমুনা উত্তর-পত্রে প্রথম উদাহরণে ঘ এর এবং দ্বিতীয় উদাহরণে গ এর নীচের খোপ দুটিতে একটি করে ক্রস-চিহ্ন (X) দেওয়া হয়েছে ।

পরবর্তী পৃষ্ঠায় ১ নং থেকে ২০ নং সমস্যাগুলি (অতীক্ষা-পত্র দ্রষ্টব্য) মনোযোগ সহকারে লক্ষ্য করো এবং নির্ভুল উত্তর বের করো । পরে উত্তর-পত্রে তোমার বেছে নেওয়া ছবি নির্দেশক অক্ষরটির (ক,খ,গ,ঘ) নীচের খোপে একটি ক্রস-চিহ্ন (X) দাও ।

কোন সমস্যা বাদ দিও না । যতক্ষণ পার চেষ্টা করো । এ পুস্তিকায় কোন দাগ দিও না । না বলা পর্যন্ত পাতা উল্টাইও না ।

উদাহরণ

#		ক	খ	গ	ঘ
১	⊥ ⊥ ⊥	⊥	⊥	⊥	⊥
২	□ □ □	□	□	□	□

নমুনা উত্তর-পত্র

#	ক	খ	গ	ঘ
১	☐	☐	☐	☒
২	☐	☐	☒	☐

টেষ্ট ঘ

নির্দেশাবলী

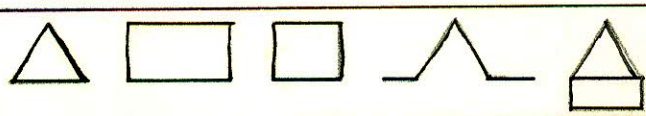
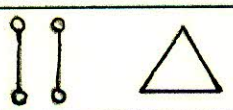
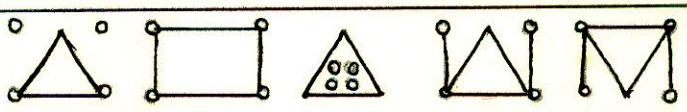
নীচে লেখা নির্দেশ গুলো পড়ে দেখ । এ টেষ্টে দেখবে প্রতি সারিতে ডান দিকে পাঁচটি (ক,খ,গ,ঘ,ঙ) এবং বামদিকে দুই বা ততোধিক ছবি আঁকা রয়েছে । ডানদিকের প্রতি সারির ছবি গুলো দুই বা ততোধিক ছবির মিলনে গঠিত । ডানদিকের প্রতি সারির ছবি গুলো দেখে এমন একটা ছবি খুঁজে বের করো যা' বামদিকের দুই বা ততোধিক ছবির মিলনে গঠিত । পরে উত্তর-পত্রে তোমার বেছে নেওয়া ছবি নির্দেশক অক্ষরটির (ক,খ,গ,ঘ,ঙ) নীচের খোপে একটি ক্রস-চিহ্ন (X) দাও ।

নীচে দুটি উদাহরণ দেওয়া আছে । প্রথম উদাহরণে বামদিকের ছবি দুটি মিলালে ডানদিকের খ ও গ অক্ষর দুটির নীচের যে কোন একটি ছবি গঠিত হতে পারে । কিন্তু ভালোভাবে লক্ষ্য করলে দেখবে গ - অক্ষরটির নীচের ছবিটি আকারে ক্ষুদ্রতর । তাই খ এর নীচের ছবিটিই সঠিক । দ্বিতীয় উদাহরণে বামদিকের তিনটি ছবি মিলালে ডানদিকের ঘ ও ঙ অক্ষর দুটির নীচের যে কোন একটি ছবি গঠিত হতে পারে । কিন্তু ভালোভাবে লক্ষ্য করলে দেখবে ঙ এর নীচের ছবিটি উল্টাভাবে রয়েছে । তাই ঘ এর নীচের ছবিটিই সঠিক । সুতরাং নীচের নমুনা উত্তর-পত্রে প্রথম উদাহরণে খ এর এবং দ্বিতীয় উদাহরণে ঘ এর নীচের খোপ দুটিতে একটি করে ক্রস-চিহ্ন (X) দেওয়া হয়েছে ।

পরবর্তী পৃষ্ঠায় ১ নং থেকে ২০ নং সমস্যা গুলো (অভীক্ষা-পত্র দ্রষ্টব্য) মনোযোগ দিয়ে দেখে প্রত্যেক ক্ষেত্রে নির্ভুল উত্তর বের করো । পরে উত্তর-পত্রে তোমার বেছে নেওয়া ছবি নির্দেশক অক্ষরটির (ক,খ,গ,ঘ,ঙ) নীচের খোপে একটি ক্রস-চিহ্ন (X) দাও ।

কোন সমস্যা বাদ দিও না । যতক্ষণ পার চেষ্টা করো । এ পুস্তিকায় কোন দাগ দিও না । না বলা পর্য্যন্ত পাতা উল্টাইও না ।

উদাহরণ

#		ক	খ	গ	ঘ	ঙ
১						
২						

নমুনা উত্তর-পত্র

#	ক	খ	গ	ঘ	ঙ
১	☐	☒	☐	☐	☐
২	☐	☐	☐	☒	☐

TEST A

Scoring Key

TEST B

Scoring Key

TEST C

Scoring Key

TEST D

Scoring Key

Test A Answer sheet

Time ..... From ..... To .....

Name ..... Male/Female ..... Class .....

Subject ..... Age ..... Name of School/College ....

Position in order of birth .....

Academic qualification and status of father .....

Academic qualification and status of mother .....

Test Score .....

#		A		B		C		D		E	
1		☐		☐		☐		☐		☐	
2		☐		☐		☐		☐		☐	
3		☐		☐		☐		☐		☐	
4		☐		☐		☐		☐		☐	
5		☐		☐		☐		☐		☐	
6		☐		☐		☐		☐		☐	
7		☐		☐		☐		☐		☐	
8		☐		☐		☐		☐		☐	
9		☐		☐		☐		☐		☐	
10		☐		☐		☐		☐		☐	
11		☐		☐		☐		☐		☐	
12		☐		☐		☐		☐		☐	
13		☐		☐		☐		☐		☐	
14		☐		☐		☐		☐		☐	
15		☐		☐		☐		☐		☐	
16		☐		☐		☐		☐		☐	
17		☐		☐		☐		☐		☐	
18		☐		☐		☐		☐		☐	
19		☐		☐		☐		☐		☐	
20		☐		☐		☐		☐		☐	

Test B

Answer sheet

Time ..... From ..... To .....

Name ..... Male/Female ... Class .....

Subject ..... Age ..... Name of School/College ...

Position in order of birth .....

Academic qualification and status of father .....

Academic qualification and status of mother .....

Test Score .....

#		A		B		C		D	
1		☐		☐		☐		☐	
2		☐		☐		☐		☐	
3		☐		☐		☐		☐	
4		☐		☐		☐		☐	
5		☐		☐		☐		☐	
6		☐		☐		☐		☐	
7		☐		☐		☐		☐	
8		☐		☐		☐		☐	
9		☐		☐		☐		☐	
10		☐		☐		☐		☐	
11		☐		☐		☐		☐	
12		☐		☐		☐		☐	
13		☐		☐		☐		☐	
14		☐		☐		☐		☐	
15		☐		☐		☐		☐	
16		☐		☐		☐		☐	
17		☐		☐		☐		☐	
18		☐		☐		☐		☐	
19		☐		☐		☐		☐	
20		☐		☐		☐		☐	
21		☐		☐		☐		☐	
22		☐		☐		☐		☐	
23		☐		☐		☐		☐	
24		☐		☐		☐		☐	
25		☐		☐		☐		☐	

Test C Answer sheet

Time ..... From ..... To .....

Name ..... Male/Female ..... Class .....

Subject ..... Age ..... Name of School/College .

Position in order of birth .....

Academic qualification and status of father .....

Academic qualification and status of mother .....

Test Score .....

#		A		B		C		D	
1		☐		☐		☐		☐	
2		☐		☐		☐		☐	
3		☐		☐		☐		☐	
4		☐		☐		☐		☐	
5		☐		☐		☐		☐	
6		☐		☐		☐		☐	
7		☐		☐		☐		☐	
8		☐		☐		☐		☐	
9		☐		☐		☐		☐	
10		☐		☐		☐		☐	
11		☐		☐		☐		☐	
12		☐		☐		☐		☐	
13		☐		☐		☐		☐	
14		☐		☐		☐		☐	
15		☐		☐		☐		☐	
16		☐		☐		☐		☐	
17		☐		☐		☐		☐	
18		☐		☐		☐		☐	
19		☐		☐		☐		☐	
20		☐		☐		☐		☐	

Test D

Answer sheet

Time ..... From ..... To .....

Name ..... Male/Female.. Class .....

Subject ..... Age ..... Name of School/College ...

Position in order of birth .....

Academic qualification and status of father .....

Academic qualification and status of mother .....

Test score .....

Rajshahi University Library
Documentation Section
Document No. D-1044
Date 9/3/88

#	A	B	C	D	E
1	☐	☐	☐	☐	☐
2	☐	☐	☐	☐	☐
3	☐	☐	☐	☐	☐
4	☐	☐	☐	☐	☐
5	☐	☐	☐	☐	☐
6	☐	☐	☐	☐	☐
7	☐	☐	☐	☐	☐
8	☐	☐	☐	☐	☐
9	☐	☐	☐	☐	☐
10	☐	☐	☐	☐	☐
11	☐	☐	☐	☐	☐
12	☐	☐	☐	☐	☐
13	☐	☐	☐	☐	☐
14	☐	☐	☐	☐	☐
15	☐	☐	☐	☐	☐
16	☐	☐	☐	☐	☐
17	☐	☐	☐	☐	☐
18	☐	☐	☐	☐	☐
19	☐	☐	☐	☐	☐
20	☐	☐	☐	☐	☐