

(i)

A GROUP TEST

of

INTELLIGENCE FOR THE DEAF.

J. M. DU TOIT.

The University of Cape Town

&&&

THESIS SUBMITTED TO THE UNIVERSITY OF CAPE TOWN
FOR THE DEGREE OF Ph.D.

September 1952.

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

SUMMARY OF

A GROUP TEST OF INTELLIGENCE FOR THE DEAF. (J.M. du Toit.)

A short survey of previous attempts at providing intelligence tests suitable for the deaf, is made, and the most important of these tests are critically examined. It is shown that the few existing group tests for the deaf are quite inadequate and unsatisfactory, a fact which might possibly account for the discrepancy in the findings of different investigators, and in the results of different tests.

The need for a better group test of intelligence for the deaf is emphasized and stated as the aim of the present investigation. For such a test to be satisfactory it would have to comply with certain principles, based on a priori considerations and on actual experience and acquaintance with the deaf. These conditions are briefly as follows:

The test should presuppose as little acquired knowledge (such as language in any form) as possible; it should be possible to give the necessary instructions by pantomime only; it should fall within the proper range of difficulty; no speed tests should be included; abundant practice material should be provided; subtests should consist of many items; the material should capture and retain the interest of deaf children.

Fifteen tests brought together for the purpose of a preliminary selection, are briefly described. The results of the application of these tests to a group of deaf pupils are considered, the merits of the different tests in practice are discussed and the reasons for discarding some of them are given.

The ten tests retained were applied to a random sample of twenty percent of all normal children of ages ten, eleven and twelve years in eight schools, together with a criterion test.

The constitution of this sample is fully discussed and the results of analyses of variance of all the different tests are presented. From these data it is concluded that the eight groups could justifiably be considered as random samples from the same population and therefore adequately representative of the total population of a thousand ten-to-twelve year old pupils in the schools concerned.

Factor analysis of the intercorrelations between these tests is performed, the influence of a first factor (a possible general factor) removed and the residual correlations shown to be insignificant. On the basis of the general factor saturation coefficients and other considerations of a practical nature six tests are chosen for inclusion in the final battery. Further detailed item-study of the chosen tests is illustrated.

The determination of the reliability of the finally constituted test battery by two different methods is described. The split-half method, applied in the case of the hearing group, yielded a reliability coefficient of .921 \pm .007; while the coefficient obtained by a repetition

of the test on a group of deaf children is given as $.969 \pm .003$.

As a measure of validity the correlation of the results of this test with those of the criterion test in the case of hearing children is presented. The method of analysis of covariance used in this connection is described. The coefficient of validity obtained in this way is given as $.802$ and the method of estimating the reliability of this coefficient by means of Fisher's Z-function is described.

The second test of validity, making ^{use} of teachers' judgments and applied in the case of deaf subjects, is then considered. The coefficients of mean square contingency obtained in five different institutions are given; these are found to vary from $.490$ to $.690$ in a case where the maximum possible is $.894$. Although these coefficients are quite satisfactory, several factors adversely influencing such coefficients are discussed.

A third test of validity, the correlation of the test results of deaf subjects with those obtained by them on a performance test, is also given. This coefficient is $.746 \pm .050$.

In order to determine the extent of possible influence of schooling and hence of command of verbal language on test results, the correlation coefficients between I.Q.'s and duration of school attendance are calculated and shown to be completely insignificant. It is concluded that the ability measured by this test is not appreciably influenced by schooling or by language mastery.

The influence of language is then considered from a different angle. Subjects were divided into two groups, those possessing naturally acquired language and those without any or with artificially acquired language only. The significance of this classification is discussed. Biserial r between these two categories is shown to be quite insignificant, which is taken as an indication of the absence of any relationship between language ability and the ability measured by this test.

In Chapter VI the distribution of raw scores of the hearing subjects is found to be skewed. The correction of this skewing by the reduction of raw scores to standard scores is described. The correlation between original and standard scores is found to be $.985$. The new distribution appears to be very much better.

Two methods of obtaining norms are then discussed.

In the first method use is made of the regression equation between the new test scores and those yielded by the criterion test, the norms of which could then be applied. Some disadvantages of this method are pointed out.

A second method, based on the frequency of occurrence of the different I.Q.'s in the general population, is described in detail and norms presented in the form of a graph. Advantages of this method are indicated, as well as certain qualifications to be taken into account.

The distribution of I.Q.'s of deaf pupils according to these norms is illustrated and found, in the case of a single school, to conform very well to a normal curve.

The average I.Q.'s of deaf children are found to vary from 98.5 (for different schools combined) to 99.96 (single school)

It is concluded that the norms, although offered as provisional and definitely dependent on those of the criterion test, appear to be reasonably useful in practice.

In Part II the actual tests are reproduced in their rough [historical] form, together with full instructions for administering and scoring and the detailed graph of norms.

C O N T E N T S .

	<u>Page.</u>
Preface	(iii)

P A R T I.

Chapter I. A SHORT SURVEY OF THE HISTORY and Present Status of Intelligence Testing and Tests for the Deaf	1
Chapter II. AIM OF THE PRESENT INVESTIGATION.	23
Chapter III. PRELIMINARY SELECTION	28
Chapter IV. FINAL SELECTION AND CONSTRUCTION of BATTERY	40
Chapter V. RELIABILITY, VALIDITY AND SOME OTHER PROBLEMS	67
Chapter VI. THE DERIVATION OF NORMS	85

P A R T II.

I. INTRODUCTION	110
II. THE TEST	113
III. INSTRUCTIONS FOR ADMINISTERING	132
IV. INSTRUCTIONS FOR SCORING	146
V. GRAPH OF NORMS	161
VI. BIBLIOGRAPHY	162

P R E F A C E .

It is a privilege to acknowledge my indebtedness to the following persons, to whom I would like to express my thanks:

My friend, Dr. K.W. Heese, who first interested me in this type of work and who put certain testing material at my disposal.

Professor J.A.J. Van Rensburg, whose interest and advice made much of this research possible.

The principals of all the schools mentioned in the text who so kindly received us, put up with the upsetting of their school routine occasioned by the testing program and provided information required, and all the assistant teachers who acted as supervisors during the taking of tests.

Special mention must be made of Sister Thomasia, principal of **The** St Vincent School for the Deaf, Johannesburg, who presented us with a complete set of the Snyders-Oomen Test, a gesture we very highly appreciated.

My wife for her assistance in the actual application of the tests, and in many other ways.

My brother, Mr. D. le R. du Toit, M.Sc., who allowed me, as member of his staff, much freedom from routine duties in order to let the project proceed. He also actively helped with the administering of the tests, read the manuscript, made many suggestions in the course of discussions, and was responsible for the partial sponsoring of the undertaking, by his school.

Our school secretary, Mrs. C. du Toit, for typing the manuscript so ably, and in such a fine spirit.

C H A P T E R I.

A SHORT SURVEY OF THE HISTORY AND PRESENT STATUS OF INTELLIGENCE TESTING AND TESTS FOR THE DEAF.

One of the first recorded attempts to estimate the mental ability of deaf children experimentally was that of Greenberger¹⁾ who presented coloured pictures and blocks to young deaf children entering school. He watched their facial expression and constructive activity. He considered sustained interest in these objects as symptomatic of fair intelligence, while complete apathy and lack of interest might indicate possible feeblemindedness.

Taylor²⁾ requested 148 deaf and a group of hearing children to write down as many words as possible in fifteen minutes. He compared the number of words and number of errors in the two groups. On the average the deaf wrote 150 words as against 153 by the hearing but made only 2.7 percent errors as against 4.3 percent by the hearing.

2...../In

-
- 1). Greenberger: "Doubtful Cases", American Annals of the Deaf, Vol. 34, 2, 1889, p. 93.
 - 2). H. Taylor: "A Spelling List", American Annals of the Deaf, Vol. 42, 1897, p. 364.

In 1899 Mott³⁾ compared eight year old deaf and hearing children by means of a number of different physical and psychological tests, notably tests of observation and memory, and found no retardation of the deaf.

A test of immediate and delayed memory was used by Smith⁴⁾ in 1903.

In 1906 MacMillan and Bruner⁵⁾ gave a number of physical and mental tests to 184 deaf children. This was probably the first attempt to apply standard psychological tests to deaf children. These tests consisted of the cancellation of A's, the judging of weights and size and the recognition of objects by touch and a test of the visual memory span for numbers. The results showed the deaf to be equal to the hearing in the judging of weights but to be retarded three to four years on the other tests. However, they were of the opinion that the deaf were simply immature and would catch up later.

With the coming of the Binet-Simon scale for the measurement of general intelligence a new era dawned in the field of intelligence testing. The preliminary stage of attempts to measure only isolated separate abilities was now followed by a period marked by attempts at devising more comprehensive systems of tests, although many investigations of more restricted type still occur.

3...../In

-
- 3). A.J. Mott: "The Ninth Year of a Deaf Child's Life", American Annals of the Deaf, Vol. 44, 1899, pp. 401 - 412.
and Vol. 45, 1900, pp. 33 - 39,
93 - 109
 - 4). J.L. Smith: "Mental Characteristics of Pupils", American Annals of the Deaf Vol. 48, 1903, pp. 248 ff.
 - 5). D.P. MacMillan and F.G. Bruner: Special Report of the Department of Child Study and Pedagogic Investigation, Chicago Public Schools, 1906
and later: "Laboratory Methods of Determining the Status of Deaf Children", Association Review, Vol. XI, No. 4, Oct. 1909, pp. 326 - 332.

In 1912 Kilpatrick⁶⁾ suggested that the Binet-Simon scale should be applied to the deaf, with adaptations if necessary. This suggestion was followed up by Pintner and Paterson⁷⁾ but, after having attempted to give the directions by speech, by finger spelling, by signs, and in writing, they concluded that the experiment was a failure and that the tests were totally unsuited for the deaf. Notwithstanding the different methods employed to convey the language content of the directions, success in each case still depended largely on language comprehension. This element could not be eliminated and caused the failure of the attempt.

A few other examples of isolated tests, mostly by the same investigators, may be mentioned here:

Learning tests of the Digit - symbol and Symbol - digit types were used with about a thousand deaf children, scoring for speed and accuracy.⁸⁾ They found retardation of two to three years and no difference between the congenitally and adventitiously deaf. The greatest value of this investigation lay in the demonstration of the possibility of using a group test with deaf subjects.

4...../From

-
- 6) W.M. Kilpatrick: "Comparative Tests", American Annals of the Deaf, Vol. 57, 1912 pp. 427 - 8.
- 7) R. Pintner and D.G. Paterson: "The Binet Scale and the Deaf Child", J.Educ. Psych., VI, 1915, pp. 201 - 10.
- 8) Pintner and Paterson: "Learning Tests with Deaf Children", Psych. Monographs, Vol. XX, 1916, No. 88.

From the results of Goddard's Formboard test, slightly modified, which was then applied⁹⁾, it appeared that the deaf were less retarded in such a manipulatory test.

In another test Pintner and Paterson found the deaf to have a poor memory span for digits¹⁰⁾. It is interesting to compare with these results those of Vertes¹¹⁾ who tested the memory of deaf and hearing children for different types of words and found the memory of the deaf for numbers the poorest, for words associated with sound images also poor, but for those with touch image better than that of hearing children. By adding a kinaesthetic sensation like tapping, the memory for digits improved to hearing standards.

Porteus¹²⁾ used his maze test on a small number of deaf children at Melbourne and concluded that the test could be used.

Amongst the above there had now been several apparently successful applications of both performance and group type of tests on a restricted scale to the deaf. The possibility of applying such tests to the deaf had been demonstrated.

5...../More

-
- 9). Pintner and Paterson: "The Formboard Ability of Young Deaf and Hearing Children", Psych. Clinic, Vol IX, Jan. 1916, pp. 234 - 7.
Also in American Annals of the Deaf, 1916, pp.184-189.
 - 10). Pintner and Paterson: "A Comparison of Deaf and Hearing Children in Visual Memory for Digits", J. Exp₁ Psych., Vol. 11, Febr. 1917, pp.76-88.
 - 11). O.J. Vertes: "Memory in Deaf Mutes", Psych. Stud. (Ranschburg - Festschrift) 1929, pp. 210 - 216 -quoted by Guilmartin: "A Summary of Psychol. Tests applied to the Deaf", The Voltz Bureau, Washington, 1932.
 - 12). S.D. Porteus: "The Measurement of the Intelligence of 653 Children examined by the Binet and Porteus Tests", J. Educ.Psych. Vol.IX, Jan.1918, pp. 13 - 31.

More comprehensive systems intended as tests of general intelligence were by now being constructed, and others have appeared since.

The most important systems of performance tests for the deaf that have been published in English are the following:

1. A Scale of Performance Tests by Pintner and Paterson¹³⁾.
2. In Britain Drever and Collins of Edinburgh constructed a series of performance tests for the deaf in 1928. A revised and better standardized edition was published in 1936¹⁴⁾.
3. A third scale that has sometimes been used is that of Grace Arthur¹⁵⁾.
4. There is also the Ontario School-Ability Examination¹⁶⁾, while Pintner's Preschool Performance Scale is probably still in the process of standardization¹⁷⁾.

We return to some of these performance tests later.

Attempts at the construction of group tests suitable for the deaf have indeed been few. The only ones of any importance in this field are the following:

1. The Pintner Nonlanguage Mental Test¹⁸⁾. This is perhaps the most ambitious and the most widely used group test for the deaf that has thus far been devised.

6...../It'

-
- 13). R. Pintner and D.G. Paterson: "A Scale of Performance Tests", Appleton, New York, 1917.
 - 14). J. Drever and M. Collins: "Performance Tests of Intelligence", Oliver and Boyd, Edinburgh, 1928; Second ed. 1936.
 - 15). G. Arthur: "A Point Scale of Performance Tests" I and II, Commonwealth Fund, New York, 1930.
 - 16). H. Amoss: "Ontario School Ability Examination", Ryerson Press, Toronto, 1936.
 - 17). Mentioned in Pintner et al.: "The Psychology of the Physically Handicapped", Crofts, New York, 1941, p. 80.
 - 18). R. Pintner: "A Nonlanguage Group Intelligence Test", J. Appl. Psych. 111, (1919), pp. 199 - 214.

It represents the first attempt at a comprehensive group test for the deaf and was used in several large scale surveys. Jeanette Reamer¹⁹⁾ applied this test together with an educational scale to 2,500 deaf children of 8 years and older, while it was used on 4,432 deaf children in a survey under the auspices of the American National Research Council²⁰⁾. This test will be discussed more fully further on.

2. The Pintner Nonlanguage Primary Mental Test²¹⁾ is intended for younger children.

3. The Chicago Nonverbal Examination²²⁾.

4. The Goodenough "Draw a Man" Test²³⁾ Although this test was not intended primarily for use with the deaf, it has quite often been applied to deaf pupils and on account of the ease with which it can be administered to them and the apparent absence of any inherent bias against the deaf, it seems to lend itself very well to such use. More about this test will be reported in the present study.

Most of the work mentioned above has been done in America. It is necessary now to refer to what has been done in some other countries.

7...../No

19). J.C. Reamer: "Mental and Educational Measurements of the Deaf", Psychol. Monogr., XXIX, 1921, No. 132.

20). Pintner, Day and Fusfeld: "A Survey of American Schools for the Deaf", National Research Council, Washington, 1928.

21). R. Pintner: "A Group Test Suitable for Younger Deaf Children", J.Educ. Psych., xxll, 1931, pp. 360 - 363.

22). A.W. Brown: "The Development and Standardization of the Chicago Non-verbal Examination", J.Appl. Psych. XXIV, 1940, pp. 36 - 47 and 122 - 129.

23). F. Goodenough: "The Measurement of Intelligence by Drawings", World Book Co., New York, 1926.

No single comprehensive scale of intelligence tests for the deaf has appeared in Germany. Work published there has often been of purely speculative nature and lacking in statistical analysis. Some of the most important publications have been the following:

One of the first tests was that of Matz²⁴⁾. He studied 2,000 examples of modelling and drawing by normal, deaf and blind children and found the deaf superior. Later G. Theil³⁵⁾ also compared large numbers of original drawings of deaf and hearing children and found the deaf superior to the hearing in perception and detail, equal in drawing from nature, but inferior in drawing from memory.

Rudolf Lindner²⁶⁾ used a great many tests of all kinds - tests of motor ability, dexterity, performance ability, drawing, memory, suggestibility, etc. This work was very extensive and suggestive of further research but formed no complete system. In some tests he found the deaf inferior, in some equal or much superior to the hearing.

Krenberger²⁷⁾ tested mainly acquired knowledge and language ability.

8...../Höfler

-
- 24). W. Matz: "Zeichen und Modellerversuch an Volksschülen, Hilfschülen, Taubstummen und Blinden", Zeitschrift für Angew. Psych., X, 1915, pp. 62 - 135.
- 25). G. Theil in Zeitschrift für Kinderforschung, 1927, 33, 138-176, referred to by Guilmartin, op.cit.
- 26). R.Lindner: "Vergleichende Intelligenzprüfungen" in Das Taubstumme Kind, Leipzig, 1925.
- 27). S. Krenberger: "Intelligenzprüfungen taubstummer Schüler", München, 1927.

Höfler²⁸⁾ tested the controlled and spontaneous abstraction of deaf children. He concluded that the amount of retardation decreased eventually in the case of controlled abstraction but not in the case of the spontaneous, from which he inferred that although the deaf were capable of rising above the plane of the concrete, they preferred to remain there as far as possible. Few subjects were used in this research. In later experiments Höfler found that command of language as well as intelligence had great influence on results of tests consisting in the arranging of series of pictures and of a criticism test, while language ability appeared to play no role in mazes and in counting the number of blocks in a diagram.

Bieri²⁹⁾ compared the development of deaf and hearing children in several respects such as concepts of space, number, objects, etc.

In Holland Herderschêe³⁰⁾ constructed a preliminary and incomplete series of tests in order to select deaf subjects fit for education. These tests were not well standardized. In some of them the average age standards appear to have decreased by seven years from 1919 to 1947 according to Oomen. Nevertheless Herderschêe was a pioneer in the sense that he was the first in Europe to realize the need for special tests for the deaf, and since then no other attempt to construct such a system of tests has been recorded there until the recent work of Oomen (see later).

9...../In

-
- 28). R. Höfler: "Über die Bedeutung der Abstraktion für die geistige Entwicklung des taubstumme Kindes", Zeitschrift für Kinderforschung XXXIII, 1927.
- 29). E. Bieri: "Ein Beitrag zur Kenntnis der geistigen Entwicklung des Taubstummen Schulkindes", 1931, Zürich and Leipzig.
- 30). Herderschêe: "Verstandsmetingen by Doofstomme Kinderen", Maandschrift v. Verloskunde, Vrouwenziekte en Kinder-geneeskunde, 1919, p. 569.

In the Rotterdamsche Inrigting voor Doofstommenonderwys Zeckel and Van der Kolk³¹⁾ recently used, amongst other tests, that of Porteus in a comparison of a hundred deaf and as many hearing subjects. They found general retardation of the deaf.

The work of Ten Have³²⁾, although fully treated statistically, concerned only 20 subjects and has nothing new to offer.

Finally there has lately appeared a very comprehensive and thorough system of performance tests for the deaf by Anna Snyders-Oomen³³⁾ further reference to which will be made later.

In the foregoing survey no attempt has been made to include every possible report about intelligence testing of deaf subjects. Reference to the work of Roth, Morrison and Bridgman with various performance tests; Peterson and Wilson, Shirley and Goodenough, and Springer with the Goodenough test; Brown, Goodlett and Greene, and Newlee; Marzi and Tavernari; for example, had to be omitted. We intended merely to give a general idea of the scope of the work that has been done and shall now consider some of the more important tests in greater detail together with the findings and conclusions reached in certain investigations with these tests.

+ + + + +

The most important performance tests for the deaf are undoubtedly those of Pintner and Paterson, of Drever and Collins,

10...../of

31). A. Zeckel and J.J. van der Kolk: "A Comparative Intelligence Test of Groups of Children Born Deaf and of Good Hearing by Means of the Porteus Test", American Annals of the Deaf, Vol. 84, 1939, pp. 114 - 123.

32). T.T. ten Have: "Over Psychologisch Onderzoek van het Dove Kind", Tydschrift voor Doofstommenonderwys, 1941.

33). A. Snyders-Oomen: "Intelligentieonderzoek van Doofstomme Kinderen", Berkhout, Nymegen, 1943.

of Grace Arthur and of Snyders - Oomen.

The first three of these have several elements in common. They are constituted as follows:

<u>Pintner-Paterson Performance Scale:</u>	<u>Arthur Point Scale:</u>	<u>Drever-Collins Performance Test:</u>
1. Mare and Foal Board	1. Knox Cube Test	1. Kohs Block Designs
2. Seguin Form Board	2. Seguin Form Board	2. Knox Cube Test
3. Five-Figure Board	3. Two-figure Board	3. Domino Test
4. Two-figure Board	4. Casuist Form Board	4. Size and Weight Test.
5. Casuist Board	5. Manikin	5. Manikin and Pro- file.
6. Manikin Test	6. Feature-profile Test.	6. Form Boards (Pintner's two- figure board and Healy's puzzle A)
7. Feature-profile Test	7. Mare and Foal Test	7. Cube Construction.
8. Ship Test	8. Healy Picture I	8. Picture Completion (Drever and Collins, and Healy's I)
9. Picture-completion (Healy)	9. Porteus Maze	
10. Knox Cube Test	10. Kohs Block Design	

Features to be noted are the heavy loading of the Pintner-Paterson Scale with tests of the form board and similar type and the fact that the Arthur Scale is identical with the Pintner Paterson in eight of its subtests, the main difference being in the method of scoring.

Pintner has consistently reported, as a result of the many investigations by himself or sponsored by him, a retardation of the intelligence of the deaf, sometimes speaking of three to four years, at others of six to seven months³⁴⁾.

11...../It

34). Pintner et al: "The Psychology of the Physically Handi-
capped", p.109 and p. 112.

It is noteworthy that whenever Pintner refers to a conflicting finding, he always queries the selection, the standardization or some other factor. In another connection Guilmartin³⁵⁾ remarks about a study by Pintner: "This study seems to indicate that even a statistician has a tendency to find what he is looking for". Drever and Collins, on the other hand, state in the 1936 edition of their book that, after having examined over a thousand cases individually by their tests, they found no significant retardation. It is not our intention to go into the merits of these different tests and the different researches. It is necessary and sufficient to state here the fact of the existence of this conflict of findings and conclusions. This controversy has served the useful purpose of stimulating interest in the problem of measuring the intelligence of the deaf, which, even so, has received much too scanty attention from psychologists. These three tests, together with the Pintner Nonlanguage scale, were the subjects of an important and careful research by MacKane, sponsored by Pintner³⁶⁾. Here, also, the main interest seems to have centred on the relative standing of the deaf and the hearing. MacKane used a group of 130 deaf children and a similar one of hearing children, carefully matched in pairs with respect to age, sex, nationality and socio - economic status of the parents. The outcome of MacKane's experiment can be summarized in his own words as follows: "The results of the Drever - Collins test "support the original conclusion of Drever that at no age level are the deaf so much as one year retarded"³⁷⁾. The other two performance scales "support in general the findings of the Drever - Collins Performance Scale"³⁸⁾. The Pintner Non-

12...../language

35). Op. cit.: p. 8.

36). K. MacKane: "A Comparison of the Intelligence of Deaf and Hearing Children", Teachers' College Contributions to Education, No. 585, Columbia University, New York, 1933.

37). Ibid. p.43.

38). Ibid. p.44.

language Test revealed at 10 and 11- year levels a marked superiority of hearing children and "re-emphasized the findings of the National Research Council by re-affirming the retardation of the deaf to the extent of two years"³⁹⁾ (In the survey of the National Council the same test was used). "It seems evident from this study that the performance scales and the nonlanguage test measure different abilities, for the study plainly shows that the same deaf children may be less than a year retarded in their responses to performance scales and yet be two years retarded in their responses to the nonlanguage test"⁴⁰⁾. It appears strange that it never seems to have occurred to MacKane or Pintner that there might be anything amiss with this nonlanguage group test which might make a close scrutiny of the test advisable.

Snyders - Oomen's system of tests is constructed on a plan similar to the Binet - Simon. It consists of a different series of tests for each age from 4 to 14 years. Some tests are common to different age levels, but a higher standard of performance is required for the higher ages. For the specially gifted there are two additional tests. The tests are also classified into four groups, the first purporting to test the perception of form and the synthesis of whole forms; the second group the realization of concrete coherence, e.g. Healy's picture, the fitting together of Pintner and Paterson's Manikin, etc; the third, abstracting ability, e.g. different tests of sorting series of pictures; the fourth, concentration of attention and immediate reproduction, e.g. Knox cubes and Drever and Collins' dominoes.

13...../On

39). Ibid: p. 43.

40). Ibid. p. 44

On account of the fewness of available subjects no attempt was made at proper standardization. Norms for different age levels were based on the results of from three to twenty subjects. Altogether only 165 subjects were used for the determination of all norms from age 3 years to age 18 years and arbitrary considerations played quite an important role in the determination of these norms. However, Snyders - Oomen, in marked contrast to many other psychologists who have investigated the problem of the intelligence of the deaf, realized that the first requisite for her task was a thorough acquaintance with the deaf and their peculiar circumstances and problems. For this reason she made a preliminary study of the deaf by careful observation in classes of an Institute for the deaf. Herein perhaps lies the main merit of her system: it is based on practical experience and apparent intimate acquaintance with the deaf. This is at once evident from her book which is notable for its penetrating insight into many different aspects and problems of the education of the deaf.

+ + + + +

For our present purpose the most important system of tests is the Nonlanguage Group Test of Pintner which therefore, requires careful consideration. It consists of the following subtests:

1. Imitation of movements.

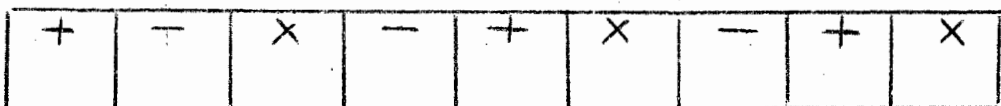
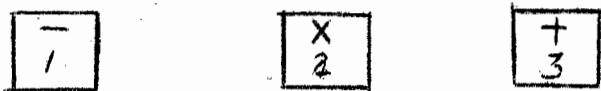
Four dots are made on the blackboard like those on the sheet in front of the child. The examiner traces with a pointer a line connecting the dots in a certain order. The child then has to join the dots on his paper in the same way. Twelve such items of four dots each are used. The test is based on the idea of the Knox cubes.

2. Easy learning test.

Three squares at the top of the page contain a symbol and a digit each. Then follow 54 squares with a symbol only, the child being required to add the appropriate digit. The time

14...../limit

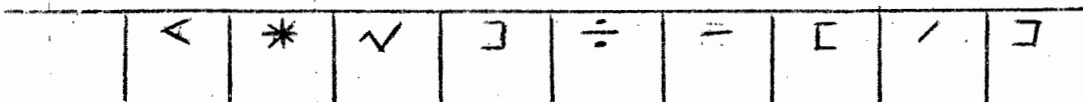
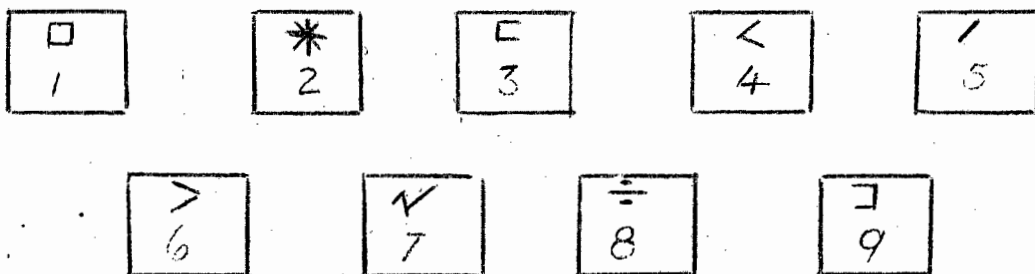
limit is one minute. Five squares with symbols are used as examples for demonstration on the blackboard.



3. Difficult learning test.

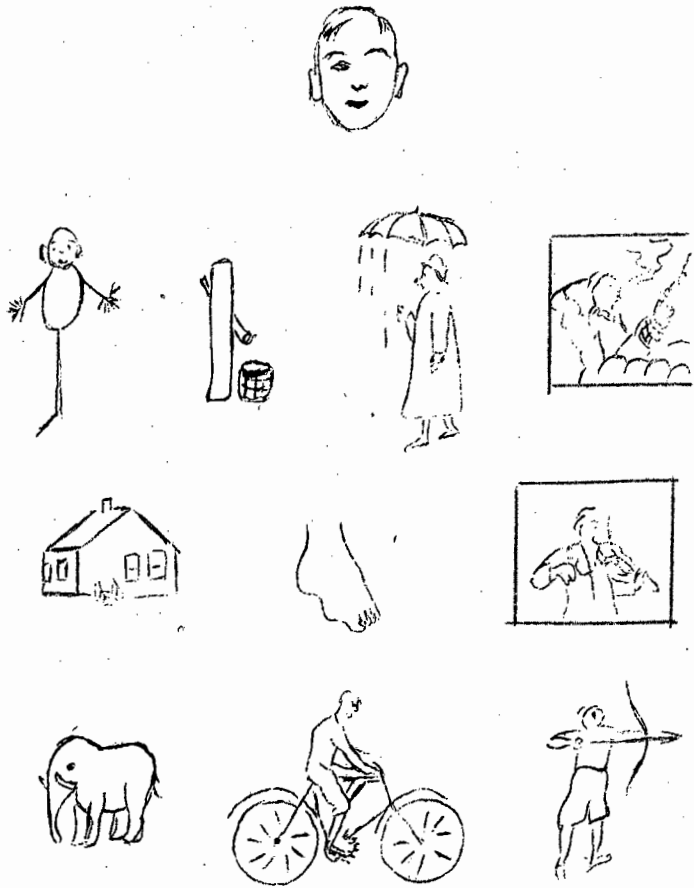
Similar to the previous but with nine symbols and digits.

Time limit: $1\frac{1}{2}$ minutes.



4. Drawing Completion.

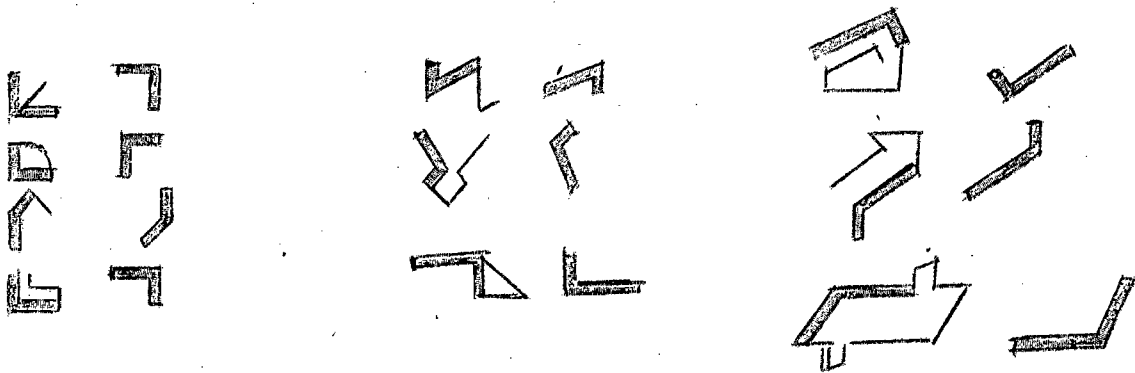
Ten drawings are given, each with some essential part missing, which has to be added by the child. One example only is demonstrated and the time limit is fixed at two minutes. The sample and the items are roughly represented below.



5. Reversed Drawings.

This test consists of a series of ten complete geometrical figures, each with two of its lines heavier than the rest; opposite each figure the heavy lines are repeated in such a way that, when completed, the figure will be upside down or reversed. The child has to add the necessary lines to complete this reversed drawing: The time limit is four minutes. Two examples are used for demonstration. The items are as follows.

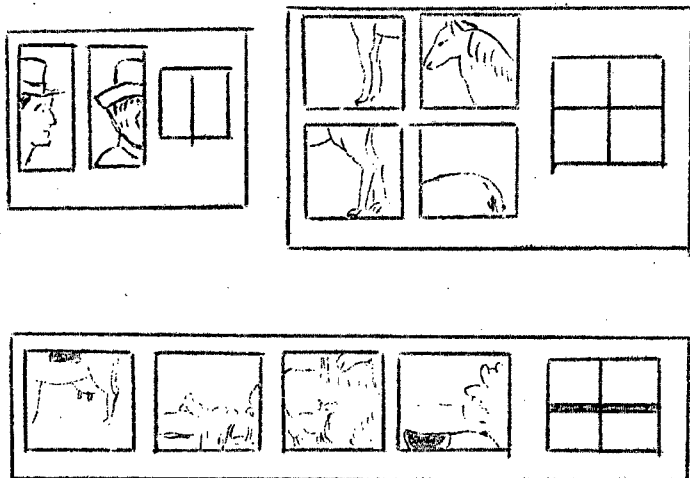
16...../figure.



6. Picture Reconstruction.

The constituent parts of a picture that has been cut up into equal squares are reproduced in haphazard order, and numbered. Alongside these there is a blank rectangle of the size of the complete picture, divided into squares of equal size to those containing the parts of the picture. The child has to indicate by numbers where each of the squares containing a part of the picture has to go in the blank in order to form the complete picture. There are five such items in the test and one example only is demonstrated. The time allowed is five minutes.

Below there are rough representations of the first three items.



These six subtests, then, constitute the Pintner Non-language Test.

In actual use in a school for the deaf we found this test to be very unsatisfactory in many respects. When a test is to be applied to two dissimilar groups, e.g. groups of different ethnological or cultural background, or with another essential difference such as radically different language ability, as is the case with the deaf and the hearing, the most elementary principle should as a matter of course be that stated by Oliver⁴¹⁾ in connection with his tests for East-African natives, viz. that only test material should be used that is equally familiar or equally unfamiliar to both groups.

Now a great many deaf children, because of various factors such as an exaggerated protective attitude of parents towards an afflicted child and the fact that institutions are often far away from the parental domicile, are sent to school much later in their lives than normal children. When they do come they have a tremendous leeway to make up. They have to start with the very elements which normal infants pick up automatically at their mothers' knees. The result is that, especially during the first years, they are educationally very far behind hearing children. Their acquaintance with letters, figures and standard arithmetical operative signs is at a given age of very much shorter duration and less intimate than that of the average hearing child of the same age, so that, both as regards easy and immediate recognition and swift and automatic reproduction of such symbols, the deaf child is relatively handicapped. While for the hearing child the

18.....\minus

41). R.A.C. Oliver: "Mental Tests for Primitive Races", Year Book of Education 1935, Univ. of London Institute of Education, and Evans Bros., London, p. 562

minus, plus and multiplication signs have, at, say, age ten years, long since attained special names and come to represent very definite concepts, the deaf child of the same age may easily still be at the stage where they are purely abstract symbols devoid of any special meaning, or otherwise where they still have to be painstakingly considered and compared. It is evident that the signs X and + are much more confusing to the latter child than to the former. Also, the mere writing of figures will come more automatically, easily and swiftly to the hearing child who has had more experience and practice with them. These considerations with reference to subtest No. 2 are seen to be of great significance when regarded in conjunction with the time limit imposed.

Similar remarks apply to subtest No. 3.

In subtest No. 4 we have several items with which the average deaf child, who is in most cases also an institutional child, probably has had neither experience nor opportunity to get acquainted. Very few such children have ever seen the type of water pump illustrated or a bow and arrow, while a great number have no knowledge of a violin, which, unless used in special acoustic classes, will in any case be an instrument without meaning for the deaf. Opportunity to observe an elephant's ear will certainly be very limited. In the case of the foot, any child might easily consider the fifth toe as being hidden behind the others and then regard the ankle as the missing part.

If one or two such items slipped in amongst twenty or thirty in a longer test, it would perhaps not be such a very serious matter. But here we find half the items of a test to be of dubious quality. Guilmartin also considers this test to be "much more of a test of environmental experience

or lack of experience".⁴²⁾

In our experience a great many children found great difficulty in understanding what was required in subtest No. 5. The level of development or intelligence at which these requirements are clearly understood is so high that a very large percentage of children fall under this line and the test then has very poor discriminatory value.

Test No. 6 we found the least useful of them all. The majority of children would choose a definite system of numbering the blank squares, apparently without reference to the pictorial material. So, e.g., one would go round in a clockwise manner, numbering the squares consecutively, another would choose an anticlockwise rotation. And yet, considering the first three items only, the clockwise method would yield a score of only one out of a possible ten, and the anticlockwise five. To assign a score of 1 to one of these performances and a score of 5 to the other is entirely misleading, as both subjects failed equally to understand the requirements and both deserved a zero score. Different individual systems were followed by others all direct evidence of complete absence of insight into the nature of the problem. And yet all these different methods yielded varying scores. In our experience this subtest was a complete and unqualified failure.

The following general criticisms also apply to this Nonlanguage Test:

1. The time limit. It is tremendously difficult to make a deaf child realize, all at the same time, that he should work as fast as possible but yet with care and accuracy; that he

20/should

42). Guilmartin, op. cit., p. 10

should start immediately at a given signal and stop at another; that he should not waste any time on erasures or over-careful drawing and attempts at perfect neatness; that if he finds he cannot solve one item he should immediately go on to the next and return to the one skipped later on if there is time; that if the point of his pencil should break, he should hold up his hand for another. And yet, if all this is not perfectly understood at the outset, a time limit becomes basically unfair, much more so in what is essentially a speed test such as subtests numbers 2 and 3 in this series.

2. The fewness of items in each subtests. When a test contains as few as ten items, chance success or failure on any one item has too much influence on the score.
3. Unsuitability of items. This has already been discussed with reference to particular subtests.
4. Inadequacy of instructions. This is one of the main shortcomings of the test, from which it would appear that the authors probably had very little experience of actual day-by-day work with the deaf. In the same connection there is also
5. The absence of practice material and
6. The absence of any method of determining whether the directions have been understood before the actual test is attempted.

It is, therefore, not surprising to find this nonlanguage test yielding consistently lower results for the deaf than any of the other types of tests (cf. MacKane's results, referred to on pages 11 and 12).

That Pintner himself was well aware of the shortcomings of his test was revealed at a conference called by the National Research Council of America after the 1928 survey

under the auspices of this Council, where he said about the test: "They are ten years old....(We see) the need' for new, longer and better nonlanguage tests for the measurement of the intelligence of the deaf".⁴³⁾ This was in 1928.

And still we find Pintner in 1941 stating in a summary of the results of all the various tests that "most weight should be laid on the results of the nonlanguage tests, because these are obtained from a very wide sampling of deaf children".⁴⁴⁾ But the widest sampling cannot make a good test out of an inherently bad one. And again, with reference to both performance and nonlanguage tests, he says: "These tests would seem to be as fair to him (the deaf child) as to hearing children"⁴⁵⁾, a statement which is open to the gravest doubts.

+ + + + +

In conclusion mention must be made of the following researches that have appeared since 1940.

W.J. Morrison, used the Ontario School Ability Examination⁴⁶⁾ on 30 deaf and 50 hearing subjects and later on 300 deaf pupils, and found the percentage of the deaf in normal and superior groups lower than for the hearing. Burchard and Myklebust⁴⁷⁾ applied the Grace Arthur test to 121 congenital and 68 adventitious deaf and found the mean I.Q.s to be 102.5 and 101.67 respectively.

22/Jane

-
- 43). Conference on the Problems of the Deaf, National Research Council, Washington, 1928; quoted by Guilmartin (op.cit.) p.13.
- 44). Pintner et al.: "The Psychology of the Physically Handicapped", p. 126.
- 45). Idem, p. 128.
- 46). W.J. Morrison: "Ontario School Ability Examination". American Annals of the Deaf, 1940, pp. 184 - 189.
- 47). Burchard and Myklebust: "Comparison of Cengenital and Adventitious Deafness...", Am. Annals of the Deaf, 1942, pp. 140 - 154.

Jane MacPherson compared 61 deaf and 66 hearing subjects by means of the Hiskey Test of Learning and found the mean I.Q.'s to be 113.87 and 101.67 respectively for the two groups⁴⁸⁾

Helen Lane⁴⁹⁾ found the mean I.Q of 242 deaf subjects on the Randall's Island Performance Series to be 102.7, and of 453 deaf subjects on the Advanced Performance Series to be 113.4.

In pointing out that the trend of recent test results seems to indicate a more normal distribution of intelligence among the deaf, Helen Lane remarks about earlier results that the instructions were sometimes not non-verbal; some test items referred to experiences not possible for the deaf; and often children were found in schools for the deaf, whose greatest affliction was mental retardation and not deafness⁵⁰⁾

Two recent researches employed Raven's Progressive Matrices, which consist of 60 items. An example of this type of test is to be found amongst the tests used in the present investigation. The first research was by Ewing and Stanton⁵¹⁾ in England on 188 subjects, the other by Oleron⁵²⁾ of the Sorbonne on 246 subjects in a Paris institution. Of these 246 only 88 were within the age groups for which there are norms for this test. The results of the two researches showed close agreement. In Oleron's words: "Results reveal marked inferiority of the deaf and slower mental development".

-
- 48). Jane MacPherson: "A Comparison of Scores of Deaf and Hearing Children on the Hiskey Test.." Unpublished Master's Thesis, Washington University, (St. Louis), 1945.
- 49). Helen Lane: "The Relation between Mental Test Scores and Future Achievement", Proceedings of the National Forum on Deafness and Speech Pathology, St Louis, 1947.
- 50). Helen Lane: "The Education of the Deaf Child" in Twentieth Century Speech and Voice Correction, ed. by E. Froeschels, Philosophical Library, New York, 1948.
- 51). Ewing and Stanton: "A Study of Children with Defective Hearing", Teacher of the Deaf, June 1943 Vol. 41, pp. 56-59.
- 52). P. Oleron: "A Study of the Intelligence of the Deaf", Am. Annals of the Deaf, Vol. 95, No. 2, March 1950 pp. 179 - 192.

C H A P T E R . II.

AIM OF THE PRESENT INVESTIGATION .

In dealing with deaf pupils, the effects of deafness and of deficient intelligence are easily confused even by experienced teachers of the deaf. It is even more difficult to form an estimate of the deaf child's intellectual abilities than it is in the case of normal children. And yet it is of the utmost importance to determine in how far the deaf pupil's educational difficulties are to be ascribed to his physical handicap and lack of language and in how far poor native intelligence may be responsible.

The complete absence of language in certain cases and the presence in others of varying degrees of natural or acquired language ability make the problem of measuring more difficult and complicated than with normal children but the need for some measuring instrument to enable us to form an objective estimate of the underlying intellectual potentialities more imperative.

While there exist today several systems of individual performance tests of intelligence suitable for the deaf, which, though not of proven value, seem to be promising and potentially useful, there is at present no satisfactory and comprehensive group test of intelligence of the more abstract paper and pencil type for them.

The purpose of the present research was to attempt to supply this need by constructing such a test battery, primarily for use with deaf subjects, but which might also be useful in all cases where language difficulties make the

application of language-type of group tests impossible or unfair, e.g. where different language groups have to be tested by a common norm, a situation frequently occurring in South Africa.

To attain this purpose several basic conditions were determined a priori by the nature of the problem and by our own experience of working with the deaf. Other of these conditions became evident in the course of preliminary try outs of many different types of tests that were being considered for inclusion.

The most important of these conditions can be briefly stated as follows:

1. The tests should presuppose as little as possible and should not in any way be dependent on schooling, acquired information such as is taught in school, degree of language acquisition or special environment. For reasons set out before, this requirement is much more essential and strict in the case of tests for the deaf than with tests for normal subjects.

No written word or letter or numerical figure should be included or required in the tests, instructions or responses.

Representations of objects should be limited to the elemental and universal, such as parts of the human body and its actions; recognition of the difference between plants and animals, flowers and leaves, etc., can also reasonably be expected. A minimum acquaintance with the most elementary and universal implements of mankind such as the axe, the hammer, chair and table, is also presupposed; but only

a very few items under this category have been included in the whole gamut of some 150 items of the final battery.

It goes without saying that in order to take the test a child must have the muscle-coordinating ability necessary to enable him to hold a pencil and to manipulate it for the purpose of producing the simple markings required. A certain amount of previous experience with paper and pencil is therefore required.

As in the case of all group-tests, the child should also, to some extent at least, be accustomed to classroom discipline and should be at a stage where he is able to keep his attention on the task for some time. For this reason a fairly long group test can only in exceptional cases succeed with children of, say, six years of age.

2. It must be possible to give the necessary instructions, without any possibility of ambiguity or vagueness, by means of pantomime and gesture only. Moreover, it must be possible to confine such instructions in their essence to a minimum. Any test requiring complicated explanations is useless. This does not mean that much time may not be required for the instruction stage, as numerous repetitions and examples are often necessary. But the essential nature of the instructions must not be involved and it must be possible to deduce what is required without any difficulty, mainly from examples. It is interesting to compare the blank expression persisting on many faces in the case of some types of tests with the quick light of understanding found in case of the successful type.

Many otherwise excellent tests had to be discarded because they could not be made to comply with this condition.

3. Tests have to come within the limits of the range of the ability of subjects intended to be tested. Every sub-test has to contain items easy enough to give the weakest subject some chance and others difficult enough to challenge the brightest.

This is, of course, nothing new, but is stated here for the purpose of pointing out that the present tests were originally devised with pupils of ages ten to twelve years only in mind. It was left for later consideration to decide whether they would be useful beyond these limits.

4. No speed tests were to be included and all restrictive time limits discarded. This point has already been discussed in the previous chapter.

This condition also serves to make the tests more suitable for use with different races. The application of certain tests to the Bantu, e.g., is sometimes criticized on the ground that the "natural tempo" of the Bantu is definitely slower than that of the Europeans on whom such tests had been standardized.

5. Abundant practice material has to be provided to get the subject acquainted with the exact nature of what is being required of him.

This practice opportunity really forms part of the instructions for each test, and has the further important function of enabling the examiner to ascertain to his complete satisfaction whether every subject has fully grasped the directions and to clear up remaining individual uncertainties.

This constitutes one of the essential features of this testing system. It also necessitates that a good deal of

time and patience should be spent on this stage in administering the tests. A poor score should be the result not of failure to understand the general nature of the task set, but merely the result of inability to solve the problem in the case of special items.

Failure to realize the importance of this condition was one of the rocks on which previous attempts at the construction of group tests for the deaf foundered.

6. Every subtest has to contain a good many items in order to provide sufficient opportunity for the child, and so that accidental success or failure on a single item will not influence the final result unduly.

As has been pointed out before, fewness of items formed one of the weaknesses of Pintner's Non-language Group Test.

7. The tests and the test material should be of such a nature and so arranged as to capture and retain the interest of the deaf subject.

Because time limits were not to be imposed, it was not necessary to arrange components of tests in strict order of increasing difficulty. The order could be somewhat varied in order to re-establish confidence and to retain interest.

It was also at an early stage decided that it would be unwise to strive after a test that could be quickly administered. This requirement of brevity would introduce a seriously hampering factor into our problem. It was realized that in order to form a reasonably reliable estimate of the general intellectual ability of a deaf subject we should have to be willing to spend ample time and patience and effort in the attempt. Ease and speed of administration, though not entirely disregarded, were therefore not accepted as among our primary conditions and goals.

CHAPTER III.

PRELIMINARY SELECTION.

Out of many tests that were considered, a series of fifteen was at first brought together, consisting of the following:

1. Symbol Sequences in pairs, some identical, others different, the identical pairs to be indicated by marking.

This test was similar to sub-test No. 5 of the American Army Group Examination Beta¹⁾, digits, however, being replaced by symbols, in conformity with the first basic condition stated in the previous chapter. Some of the first items are given below.

X ~ O _ _ _ X ~ O

O < ^ _ _ _ O < v

They gradually increased in length to something like

> < □ < X ~ + v ^ O _ _ > < □ < X ~ + ^ v O

2. Completion of Symbol Series. This was something of the nature of subtest No. 3 of the Army Beta, and similar

29...../in

1). Yoakum and Yerkes: "Mental Tests in the American Army", London, 1920.

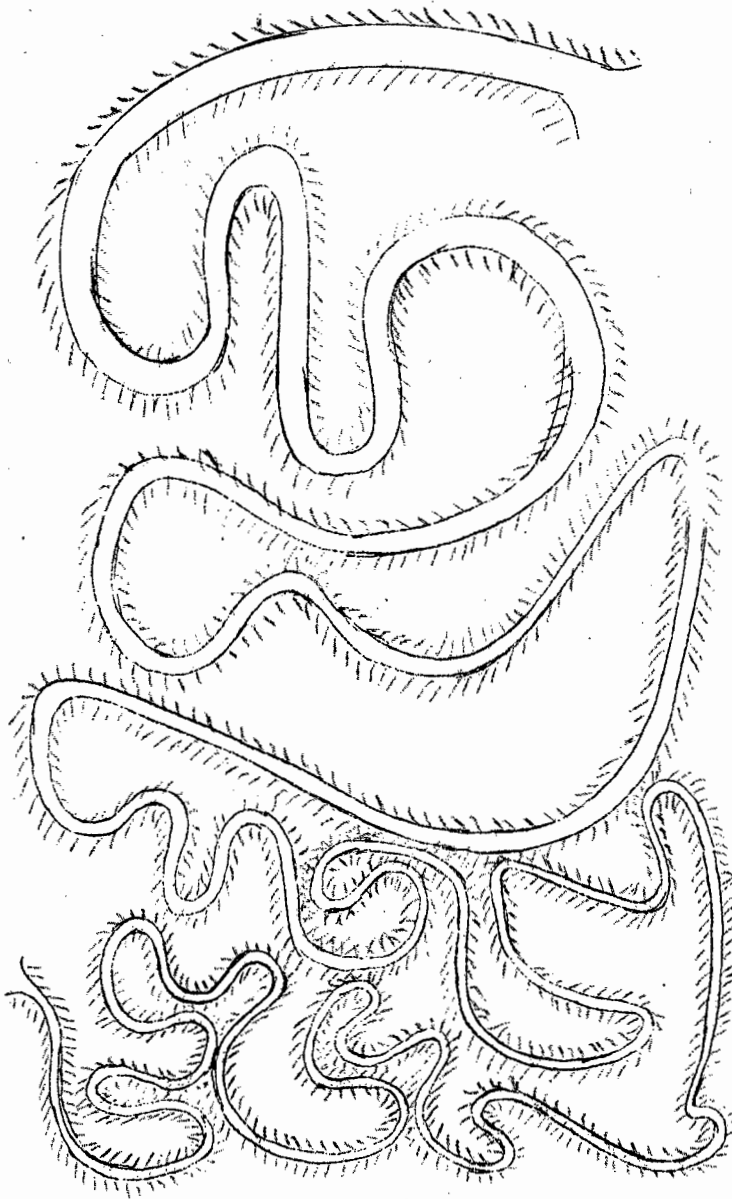
in principal to No. 3 of the Kingsbury Primary Test²⁾, some of the items of which were incorporated. The test will be reproduced in full later.

3. Classification. In a series of five or six representations of objects or abstract figures which could be classified as belonging together with respect to some single attribute they have in common, one was included that did not fit into this category and had to be detected.

4. Motor Co-ordination. A serpentine path formed by two lines gradually approaching each other and so narrowing the path, had to be traversed by drawing a line in pencil between them, an effort being made to avoid touching the bounding lines. Every contact with them was counted as a minus mark. See figure on next page.

30...../fig.

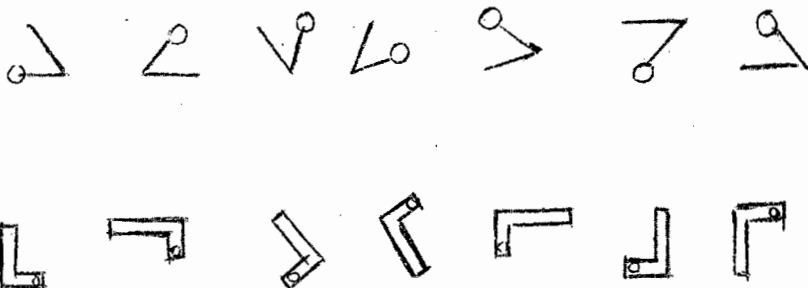
2). Kingsbury Primary Group Intelligence Scale, Public School Publishing Co.



5. Matrix Completion. Sixteen matrices, each with two rectangular parts missing, had to be completed by selecting the appropriate rectangles to supply the missing parts of the matrix pattern from a group of six provided for each matrix.

This test, adapted from one used by Professor Reyburn of Cape Town University³⁾, and derived from the Raven Matrix-tests, will be reproduced in full at a later stage.

6. Figure Manipulation. This test was derived from Thurstone's S-factor tests.⁴⁾ There were twenty rows, each consisting of seven identical geometrical figures in different positions, a few in each row, however, being mirror images of the first. The other figures in the row could be made to coincide with the first by merely rotating them in the plane of the paper. The mirror image-figures could, of course, only be brought to coincidence with the first figure by being lifted out of the plane and reversed. These were the figures to be identified and marked. Examples are given below:



7. The Goodenough Draw-a-man test. The children were requested to draw a man as best they could. The drawings were marked with disregard for artistic merit, marks being allotted for the presence of certain details, the proportion of parts, motor co-ordination, etc., according to a standardized schedule⁵⁾. 32...../ 8. F - test.

3). Obtained through the courtesy of Prof. J.A.J. Van Rensburg, of the University of Stellenbosch.

4). Obtained from Thurstone through the courtesy of Dr.K.W. Heese, formerly of Stellenbosch. Later published in: L.L. and T.T.G.Thurstone: Factorial Studies of Intelligence (Psychometric Monographs No.2.)University of Chicago, 1941.

5). F. Goodenough: "Measurement of Intelligence by Drawings", New York, 1926.

8. F - test. This test, based on one of Cattell's f-factor ("fluency of association") tests⁶⁾, consisted of several sheets, each containing six identical reproductions of some simple figure consisting of one or two lines only. Out of these, by the necessary additions, representations of different objects had to be made.

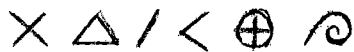
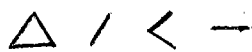
For example, the first page contained the figure



six times, from which representations of objects like a chair, a table, etc., could be made. The purpose was of course to see how many different things could be thought of with the given lines as germ idea.

9. Figure Recall. A series of geometrical figures of increasing complexity, each exhibited for a short period, had to be reproduced from immediate memory.

10. Memory Series. This test was similar to the previous one except that instead of one or two large figures each card contained a series of simple small geometrical figures which had to be reproduced.



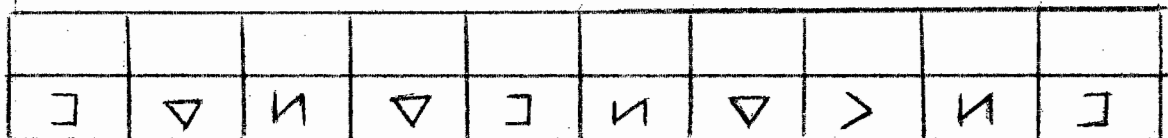
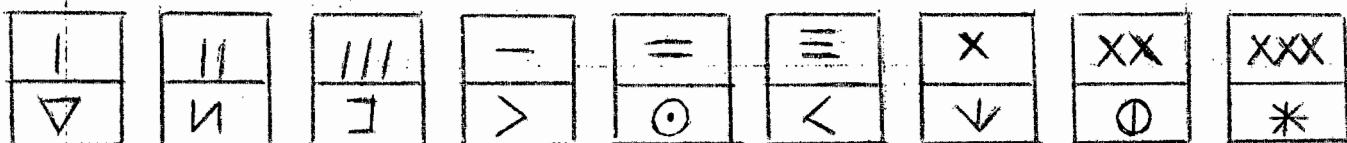
11. A Maze Test. The maze was one used by Reyburn⁷⁾,
33...../consisting

6). Cattell: "A Guide to Mental Testing" - Univ. of London Press, 1936.

7). Obtained through the courtesy of Prof. Van Rensburg.

consisting of 30 smaller figures all connected together to form one continuous maze.

12. Associated Symbols. This was a test of the well-known digit-symbol type, as used in the Army Beta scale. Two tests of this type were included by Pintner in his Non-language Group Test⁸⁾. In the present test the digits of the usual type of test were replaced by a second set of symbols in accordance with our stated principles, as in the following example:

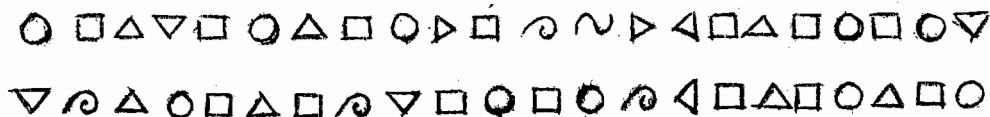


etc.

13. Identical Pairs. In each row of five similar figures the only two identical ones had to be selected. This test is similar to Thurstone's Faces-test⁹⁾ from which some of the items were derived.

14. Sequence of Three Symbols. In a lengthy unbroken series of small figures every occurrence of the group

△ □ ○ in this exact sequence had to be encircled.



34...../Cancellation.

8). See pp.13. 14, in Chapter I.

9). Thurstone, op. cit.

TABLE I. INTERCORRELATIONS BETWEEN FIFTEEN PRELIMINARY TESTS.

Test No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15.
1		.31	.01	-.16	.26	.24	.12	.15	.03	-.07	.23	.30	.12	.22	.15
2			.52	.14	.55	.28	.54	.34	.48	.25	.36	.47	.56	.61	.57
3				.29	.33	.21	.58	.41	.42	.10	.47	.51	.44	.22	.38
4					.15	.15	.29	.20	.34	.23	.42	.25	.35	.20	.54
5						.32	.56	.22	.56	.17	.49	.49	.05	.21	.30
6							.51	.22	.40	.08	.48	.28	.11	.33	.17
7								.54	.62	.24	.62	.51	.06	.12	.40
8									.44	-.03	.46	.13	.25	.24	.12
9										.48	.57	.34	.18	.37	.47
10											.29	.47	.38	.50	.28
11												.46	.21	.27	.70
12													.31	.09	.14
13														.48	.38
14															.32
15															

36...../The

- 35 -

The first factor loadings yielded by the factorial analysis are given in Table 2 and the residual correlations, after removal of the first factor effect in Table 3. (The first factor we here regard as the general factor at the measurement of which we are aiming).

TABLE 2. FIRST FACTOR LOADINGS OF 15 PRELIMINARY TESTS.

Test No.	Saturation Coefficient.	Test No.	Saturation Coefficient.	Test No.	Saturation Coefficient.
1	.232	6	.472	11	.783
2	.774	7	.738	12	.603
3	.623	8	.460	13	.485
4	.420	9	.738	14	.525
5	.590	10	.420	15	.626

37...../table 3

TABLE 3. RESIDUAL INTERCORRELATIONS AFTER REMOVAL OF FIRST FACTOR.

Test No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
1		.13	-.13	-.26	.12	.13	.05	.04	-.14	-.17	.05	.16	.01	.10	.00	
2			.04	-.18	.09	-.08	-.03	-.02	-.09	-.07	-.25	.00	.19	.20	.09	
3				.03	-.03	-.08	.12	.12	-.04	-.16	-.02	.13	.14	-.11	-.01	
4					-.10	-.05	-.02	.01	.03	.05	.09	.00	.15	-.02	.28	
5						.04	.12	-.05	.13	-.08	.03	.13	-.24	-.10	-.07	
6							.16	.00	.05	-.12	.11	.00	-.12	.08	-.13	
7								.20	.07	-.07	.04	.06	-.30	-.27	-.06	
8									.10	-.22	.10	-.15	.03	.00	-.17	
9										.17	-.01	-.10	-.18	-.01	.01	
10											-.04	.22	.18	.28	.02	
11												-.01	-.17	-.14	.21	
12													.02	-.23	-.24	
13														.22	.08	
14																-.01
15																

(Value of r required for significance at 1% and 5% levels (40 subjects): .403 and .312 respectively)

38...../The

The method of analysis will be more fully treated at a later stage.¹⁰⁾

Much weight could not be attached to this analysis, since it was based on so few subjects. The first factor saturation coefficients were merely used as additional indication of the value of tests. Decision to discard or retain tests at this stage was however mainly based on observation during the taking of the tests and critical examination of the examination papers.

In the light of experience and results obtained the following tests were now eliminated:

No. 4. The Serpentine. This appeared to be a very difficult test to mark reliably and quickly. A great deal of the success or failure depended on external factors, such as the sharpness or bluntness of the pencil point and the nature of the surface under the paper.

No. 6. Figure Manipulation, which appeared to be too difficult. It was found to fail for much the same reasons as Pintner's reversed drawings.¹¹⁾

No. 8. F - test. In this case it was found impossible to attain an objective system of marking. In many cases it was difficult to decide whether any definite object was intended. This made marking much too arbitrary. In some cases the testee would write the name of the object he had drawn. As another, with perhaps an equally vague drawing, was unable to name his effort in writing,

39...../the

10). See Chapter IV.

11). See Chapter I, p. 15

the language factor, which it was hoped to eliminate, was re-introduced.

No. 10. Memory Series. This was too similar to No. 9 and apparently not as good. (Compare Saturation Coefficients).

No. 15. Cancellation. With the time limit removed, this test lost most of its discriminatory value, the distribution becoming very unsatisfactory. Most of the scores were bunched together high up on the scale. No. 1, was retained notwithstanding its low saturation coefficient because in practice it seemed to show some promise at this stage.

The tests retained were also carefully scrutinized and in some cases radically amended by the removal of some items, the introduction of new items, rearrangement of items and alterations in the method of presentation and in the practice material.

CHAPTER IV.

FINAL SELECTION AND CONSTRUCTION OF BATTERY.

The tests retained after the preliminary selection consisted of the following ten:

1. Pintner's Four Dots¹⁾, which was now brought into the series.
2. Completion of Symbol Series, as before.
3. Classification, as before.
4. Matrices, formerly No. 5.
5. Goodenough Test, formerly No. 7.
6. Figure Recall, formerly No. 9.
7. The Maze, formerly No. 11.
8. Identical Pairs, formerly No. 13.
9. Symbol Sequences in Pairs, formerly No. 1.
10. Sequence of Three Symbols, formerly No. 14.

In this chapter these new numbers will be used to designate the different tests.

These ten tests were applied to a random sample of twenty percent of all pupils of ten, eleven and twelve years of age in eight schools for normal children,²⁾ as shown in Table 4.

40...../Table 4.

-
- 1). See Chapter I, p.
 - 2). By kind permission of the Cape Provincial Department of Education and the principals of the schools concerned.

Table 4. SCHOOLS FROM WHICH STANDARDIZING SAMPLE WAS DRAWN, AND NUMBER OF PUPILS USED IN EACH CASE.

SCHOOLS	NUMBERS.
Goudini High School, Rawsonville	19
Wolseley Secondary School	16
Tulbagh High School	20
Ceres Primary School	28
Worcester Girls' High School	34
Worcester Boys' High School	36
De Doorns Secondary School	30
Villiersdorp Primary School	17
Total ..	200

In addition to these tests, the South-African Group Test of Intelligence³⁾, form A, was also applied to the same sample. This test, very well standardized on nearly 17,000 children in the Union of South Africa, was chosen as a reliable criterion.

The Sample.

To obtain a random sample, a complete list of all pupils of the 10, 11 and 12 year age groups was

42...../obtained.

3). R.W. Wilcocks: Die Suidafrikaanse Groep-Verstandstoets, Univ. of Stellenbosch, 1931.

and R.W. Wilcocks: "On the Distribution and Growth of Intelligence", Journal of General Psych., 1932 Vol. VI, No. 2.

obtained from each school. The selection of twenty per cent of such pupils was then made in each case by means Fisher's tables of random numbers⁴⁾. A random sample of 200 such pupils was thus obtained, representative of a population of 1,000 ten to twelve - year old European school-going children. To test this claim that the eight distinct groups could be regarded as constituting a random sample from the same population, we had to determine whether there were any significant differences between the results of the different groups. This required an analysis of the variance of the sample, a method developed by Fisher⁵⁾, and adapted for use in educational research by Lindquist⁶⁾.

The variance of a sample consisting of a number of different groups can be shown to contain two components, viz. the average variance within the groups and the variance of the group means themselves. As variance can be defined as the sum of the squared deviations from the mean,
degrees of freedom

the above statement can be put in the form (omitting the degrees of freedom for the time being):

Total sum of squared deviations of all scores from the general mean = sum of squared deviations of every score from the mean of its own group + sum of squared deviations of the group means from the general mean.

43...../For

4) R.A. Fisher & Yates: "Statistical Tables for Biological, Agricultural and Medical Research", Edinburgh.

5) R.A. Fisher. "Statistical Methods for Research Workers", Edinburgh, 1925 - 1948.

6) E. F. Lindquist: "Statistical Analysis in Educational Research", Houghton Mifflin Co, Boston, 1940.

For this we use the following abbreviated notation:

$$SS_{\text{total}} = SS_{\text{within groups}} + SS_{\text{between groups}}$$

or, transposing,

$$SS_{\text{within groups}} = SS_{\text{total}} - SS_{\text{between groups}}$$

Now, if the different groups constituting a large sample are random samples from the same population, two independent estimates could be made of the population variance, one being based on the first of the above-mentioned components, i.e., on the average variance within groups only, the other on the second component only, i.e., on the variance of the group means, or, as we expressed it above, the variance between groups; and these two estimates would differ by chance only. This can be illustrated diagrammatically as follows:

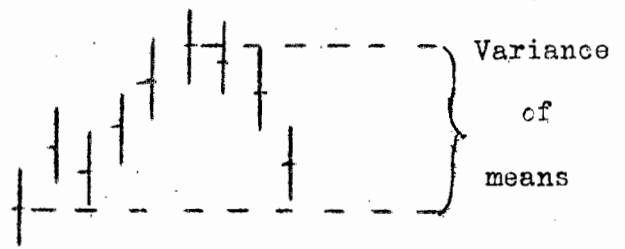
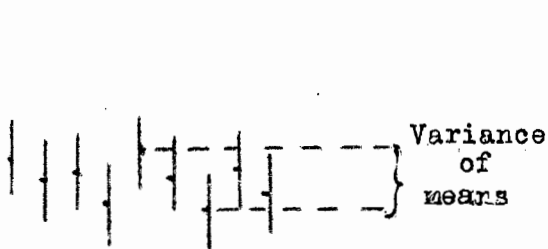


Fig. IV. 1. Variances of different Groups, their Means differing by Chance only.

Fig. IV. 2. Variances of different Groups, their Means differing significantly.

It is evident that in fig. IV. 2, where there is a significant difference between the groups, the variance between the means is greater than the average variance within groups. In the case of fig. IV. 1 we may reasonably expect the means themselves to show about as much variance as the average variance within the groups.

If the ratio between the two estimates, that based on the variance within groups and that based on the variance between groups, prove significant, we would have to conclude that the variance of the different group means could not be due to chance only and that there must therefore exist systematic differences between the groups used, which could then not be considered random samples from the same population of ten- to - twelve - year old children.

Our immediate purpose is then to determine the ratio of the variance within schools to the variance between schools and to test this ratio for significance.

The total "sum of squares" is first found, then the "sum of squares" between schools, and from these two, by subtraction, the "sum of squares" within schools.

For purposes of calculation use can be made of the following alternative expressions for sums of squares in each case:

1.
$$SS_{total} = \sum d'^2$$
 where d' represents the deviation of every individual score from the general mean.

$$SS_{total} = i^2 \left[\sum fd''^2 - \frac{(\sum fd'')^2}{N} \right]$$
 where d'' is the deviation from an arbitrary

mean, and i the size of the class intervals.

$$2. SS_{\text{between schools}} = \sum n_p (GM - M_p)^2$$

where GM denotes General Mean, M_p the mean of a group,

n_p the number of cases in a group.

or

$$= \frac{T_1^2}{n_1} + \frac{T_2^2}{n_2} + \frac{T_3^2}{n_3} + \dots + \frac{T_g^2}{n_g} - \frac{GT^2}{N}$$

where T_i denotes the total of the first group, GT the Grand Total, and N the total number of cases.

The second form in each case is preferred, to facilitate calculations.

In Table 5 is given the information necessary to calculate the variances required for some of the different tests separately. (The data are only given for the tests that were eventually accepted for inclusion in our battery).

Table 5 .

DATA FOR THE CALCULATION OF VARIANCES.

TEST. No.	CLASS Interval.	$\sum fd^2$	$\frac{(\sum fd)^2}{N}$ (N = 200)	Totals in different Schools								G.T.
				$n \frac{1}{= 19}$	$n \frac{2}{= 16}$	$n \frac{3}{= 20}$	$n \frac{4}{= 28}$	$n \frac{5}{= 34}$	$n \frac{6}{= 36}$	$n \frac{7}{= 30}$	$n \frac{8}{= 17}$	
Criterion	1	5161	136.13	211	161	207	324	361	419	327	225	2235
1	1	1102	3.92	127	100	132	205	236	254	175	143	1372
2	1	4308	19.22	338	302	383	551	679	728	588	369	3938
3	1	1973	351.125	318	273	333	477	592	616	578	255	3442
4	1	2493	0.125	134	121	178	267	314	355	268	158	1795
6	1	5959	39.60	215	201	214	387	397	462	377	258	2511
8	1	6226	0.50	284	266	344	523	669	640	551	333	3610
(Combination of nos. 1, 2 (3, 4, 6 & 8) ($\frac{1}{2}$ 5)	1	3766	8.82	307	272	344	529	634	672	544	340	3642

Table 6 . SUMS OF SQUARES, VARIANCES AND RATIOS BETWEEN VARIANCES FOR DIFFERENT TESTS.

Test No.	Sums of Squares.		Degrees of Freedom.	Variance.	F (Ratio of Variance between to Variance within Schools.)
Criterion	Total	5024.87	199	Not required	
	Between Schools	130.67	7	18.67	$\frac{18.67}{25.49} = .73$
	Within Schools	4894.20	192	25.49	
1	Total	1098.08	199	-	
	Between Schools	88.00	7	12.57	2.39
	Within Schools	1010.08	192	5.26	
2	Total	4288.78	199	-	
	Between Schools	167.30	7	23.9	1.11
	Within Schools	4121.48	192	21.5	
3	Total	1621.175	199	-	
	Between Schools	223.38	7	31.91	4.38
	Within Schools	1397.795	192	7.28	
4	Total	2492.88	199	-	
	Between Schools	143.46	7	20.49	1.67
	Within Schools	2349.48	192	12.24	
6	Total	5919.40	199	-	
	Between Schools	288.72	7	41.25	1.41
	Within Schools	5630.68	192	29.33	
8	Total	6225.50	199	-	
	Between Schools	377.04	7	53.86	1.77
	Within Schools	5848.46	192	30.46	
Combina- tion of the above	Total	3757.18	199	-	
	Between Schools	205.54	7	29.36	1.59
	Within Schools	3551.64	192	18.50	

According to Snedecor's⁸⁾ table, reproduced as Table 4 by Lindquist⁹⁾, the value of F , for 7 and 192 degrees of freedom which might be expected to be exceeded in less than five percent of cases by chance, if there exists no real difference, is 3.252. In other words, an F resulting from chance only will be less than 3.252 in 95 cases out of 100. Similarly, and from the same table, it is seen that in 99 cases out of 100 an F resulting from chance only will be less than 5.704. To be significant of a real difference between schools F must therefore exceed 3.252 at the 5 percent or 5.704 at the one percent level of significance.

According to this test all the F -values in Table 6 are seen to be insignificant, except in the case of test number 3, where the F is significant at the 5 percent but not at the one percent level.

We may therefore safely conclude that the difference in school means may well be the result of chance fluctuations in random sampling alone and we are justified in considering the eight groups, as far as these tests are concerned, as random samples from the same population and therefore adequately representative of the total population of 1,000 ten-to-twelve year old school children from which they were drawn.

+ + + + +

The correlation coefficients of each of the ten tests with the criterion test were calculated and are given in Table 7

These coefficients already give some indications as to the suitability of the several tests.

50.... /Table 7

8) G.W. Snedecor, "Statistical Methods", Collegiate Press, Ames, Iowa.

9) Lindquist, Op. cit.

Table 7 . COEFFICIENTS OF CORRELATION WITH THE S.A. GROUP TEST.

Test No.	Coefficient of Correlation.	
1	.495	Values of coefficient required for significance at the 5 percent and 1 percent levels = .139 and .182 respectively.
2	.662	
3	.525	
4	.675	
5	.432	
6	.557	
7	.417	
8	.549	
9	.377	
10	.444	

51...../Table 8

Table 8 . INTERCORRELATIONS OF TESTS.

Tests	C ^x	1	2	3	4	5	6	7	8	9	10
C		.495	.662	.525	.675	.432	.557	.417	.549	.377	.444
1	.495		.395	.416	.448	.173	.453	.391	.416	.310	.308
2	.662	.395		.571	.664	.335	.585	.452	.610	.402	.405
3	.525	.416	.571		.498	.234	.421	.408	.492	.380	.434
4	.675	.448	.664	.498		.471	.589	.490	.641	.430	.504
5	.432	.173	.335	.234	.471		.533	.192	.310	.193	.300
6	.557	.453	.585	.421	.589	.533		.491	.538	.339	.434
7	.417	.391	.452	.408	.490	.192	.491		.397	.389	.413
8	.549	.416	.610	.492	.641	.310	.538	.397		.548	.484
9	.377	.310	.402	.380	.430	.193	.339	.389	.548		.420
10	.444	.308	.405	.434	.504	.300	.434	.413	.484	.420	

52...../Factor

- 51 -

^x C denotes the S.A. Group test, here treated on equal footing with the others

Factor Analysis.

All the intercorrelations of these tests are given in Table 8 .

We now wished to determine whether the presence of a "general" factor common to all these tests could, with or without further "group" factors, account for these intercorrelations. This common factor - if its presence could be established - we took to be the ability we were attempting to measure, and which, in this research, is denoted by the term "intelligence"¹⁰⁾. Assuming such a factor to be present in our tests, we would require a measure of the degree of correspondence of each test with the estimated measurement for this factor in order to decide which tests measure best that which we attempt to measure. Such a correlation of a test with a hypothetical factor is often called the "factor loading" or "factor saturation" and indicates "the degree to which each test - performance appears to depend on the fundamental ability or 'factor' supposed to influence it."¹¹⁾ Furthermore, the possibility of "group" factors, restricted to some tests only, being partly responsible for high intercorrelations had to be taken into account. Tests with high group factor loadings would have to be excluded as we wished to avoid as far as possible our results being influenced by restricted or specific abilities.

Reduced to its simplest form the theory involved, as far as we require it for our present purpose, amounts to the following:

53...../The

-
- 10). We are not here concerned with the nature of this factor, nor with the problem as to whether this factor itself might be a complex of other sub-factors into which it might possibly be analyzed by the use of more tests yielding a much larger matrix.
- 11). Cyril Burt: "The Factors of the Mind", Univ. of London Press, 1940. P. 6.

The correlation between two tests may be regarded as the result of the influence on the performance in both cases of one or more identical mental abilities. A fictitious measure of this hypothetical ability may be termed a factor. If the saturations or loadings of Tests A and B with the first common factor, i.e., the correlation between these tests and the factor, are given by a and b , then the presence of this common factor would cause a correlation between tests A and B equal to the product of these saturations, i.e., $a \times b$. If the total correlation between A and B amounts to c , then a part of it, equal to ab , is due to the presence of this first common factor. After removal of this part, any residual correlation, given by $c - ab$, if significant, must be caused by one or more further factors; and so on.

Given a matrix of intercorrelations between several tests, we attempt by means of factor analysis to obtain the saturations with one or more common factors that could give rise to such intercorrelations. Under common factors we here include factors common to all the tests concerned - so-called general factors - and also those common to some of them only - known as group factors. We seek after the simplest structure, that is, the smallest number of common factors that could possibly account for the system of intercorrelations.

There are several different methods of analysis available which lead to essentially the same results, as pointed out by Burt¹²⁾, whose method we here follow. As

54...../we

12). Burt, op. cit.

we are mainly interested in the first general factor, and thereafter only in the possible significance of residuals, we employ the weighted summation method of single factor analysis of which Burt says: "Like Thurstone's method, it yields a factor - pattern of minimal rank; like Hotelling's, it yields the best possible fit. It would thus seem to combine the merits of both"¹³⁾

To obtain the best weights the simple summation method is first employed. The formula for the saturation coefficient by this method is given as

$$r = \frac{\sum_i r_{ai}}{\sqrt{\sum_i \sum_j r_{ji}}}$$

where the subscripts a and g refer to the test concerned and the general factor; $\sum_i r_{ai}$ represents the sum of the correlations of this test with every other test in the matrix and $\sum_i \sum_j r_{ji}$ the grand total of all the intercorrelations.

The original matrix, Table 8, is rearranged according to the totals of the columns, in Table 9. The same table also contains the further calculations. In the line indicated by "Self-correlations, first estimate", provisional estimates are given for the blank diagonal cells in the table. These are added to the totals immediately above them to give the first completed totals for the columns in the next line. These completed totals are added and the square-root of their sum, here equal to 7.34, used as divisor. The quotients obtained by dividing this into the completed totals for each column represent the first estimates of the saturation coefficients, in the next line. These coefficients are squared, and the squared values, or self-correlations, are compared with the estimated self correlations three lines higher up. As they do not agree too well, the whole process is repeated with adjusted estimates for the self-correlations. The coefficients yielded by the third estimated self-correlations have squares that agree sufficiently well with the estimated values. These coefficients might themselves now be accepted as the saturations required, but in the method of weighted summation they are used as weights for the next stage.

55...../Table 9

13). Burt, Op. cit., p.486

TABLE 9 INTERCORRELATIONS REARRANGED ACCORDING TO TOTALS AND CALCULATION OF PRELIMINARY SATURATION

The saturation coefficients obtained are now used as weights for the weighted summation method. The formula in this case is

$$r_{og} = \frac{\sum_i r_{ig} r_{oi}}{\sum_i \sum_j r_{ig} r_{ji} \div \sum_i r_{ig}}$$

In Table 10 are given the products obtained by multiplying each row of the correlation table, as used in Table 9 - but with the self-correlations inserted in the diagonal blank cells - by the weights. The grand total of all the columns is divided by the total of the weights, yielding a quotient of 5.042, which in turn is used as divisor to divide into the totals of the columns individually. The quotients so obtained are the revised estimates for the saturation coefficients. As these conform closely enough to the weights - the only differences appearing in the third decimal figure except in the case of the last test - we did not find it necessary for our purpose to repeat the process, using these saturation coefficients in their turn as weights, but accepted them as our saturation coefficients. The total of the squares of these coefficients, 5.040, is also very nearly equal to the quotient obtained before by dividing the grand total of the columns by the total of the weights (viz. 5.042) so that no further adjustment in this respect is necessary.

A hierarchy of theoretical correlations resulting from this first factor is now calculated by multiplying each of the saturation coefficients by every other one. These general factor correlations are entered in the top right hand half of Table 11, underlined in red. Above each is the corresponding original observed correlation, and below the residual obtained by subtraction of the Theoretical general factor correlation from the observed correlation. These residuals are repeated in the lower left half of the table so as to stand out clearly.

Table 10. WEIGHTED INTERCORRELATIONS AND CALCULATION OF SATURATION COEFFICIENTS.

Weights	T e s t s											Total.
	4	C	2	8	6	3	10	7	1	9	5	
.8312	.5742	.5610	.5520	.5328	.4896	.4139	.4189	.4072	.3724	.3575	.3914	
.7830	.5285	.4799	.5184	.4299	.4362	.4111	.3477	.3265	.3877	.2952	.3383	
.7738	.5139	.5123	.4633	.4720	.4527	.4418	.3134	.3497	.3056	.3111	.2592	
.7572	.4854	.4158	.4618	.4341	.4074	.3726	.3664	.3006	.3150	.4150	.2347	
.7496	.4416	.4177	.4386	.4033	.4213	.3156	.3254	.3681	.3396	.2542	.3996	
.6550	.3261	.3439	.3740	.3222	.2757	.2809	.2843	.2673	.2727	.2489	.1532	
.6166	.3108	.2738	.2498	.2984	.2676	.2676	.2344	.2547	.1900	.2589	.1849	
.5993	.2936	.2499	.2708	.2379	.2943	.2445	.2475	.2152	.2343	.2331	.1150	
.5613	.2515	.2778	.2217	.2335	.2543	.2335	.1729	.2195	.1768	.1740	.0971	
.5586	.2402	.2106	.2246	.3061	.1893	.2122	.2346	.2173	.1732	.1744	.1078	
.4613	.2173	.1993	.1545	.1430	.2459	.1079	.1384	.0886	.0798	.0890	.0982	
Totals: 7.3469	4.1831	3.9420	3.9295	3.8132	3.7343	3.3016	3.0839	3.0147	2.8469	2.8113	2.3794	37.0399 ÷ 7.3469) = 5.042
Saturation Coef- ficients	.830	.782	.779	.756	.741	.655	.612	.598	.565	.558	.472	
Squares	.6882	.6113	.6074	.5720	.5485	.4287	.3740	.3575	.3188	.3109	.2227	5.040

58...../According

According to Lindquist's ¹⁴⁾ calculation based on Fischer's procedure for determining whether a correlation coefficient obtained from a small random sample is significant, we require for significance in the case of a sample of 200 a coefficient of .139 at the 5 percent level and of .182 at the one percent level. From this it follows that only two residuals are of any possible importance, viz.

$r_{8,9} = .126$, which does not quite reach the 5 percent level, and

$r_{5,6} = .184$, which just reaches the one percent level. However, this cannot be regarded as significant as it is merely what is to be expected as a result of chance, as pointed out by Burt¹⁵⁾. (Nevertheless, one test of each pair has been excluded from the final battery, although also for other reasons.)

The saturations of our tests with a first, general factor, the measurement of which is the goal of our research, can therefore account for practically all the intercorrelations between our tests. There appears to be no indication of further or group factors influencing our tests appreciably. It is therefore not necessary to discard tests on account of high intercorrelations as these have been shown to be a consequence of high saturations with the same general factor.

59...../Table 11.

14.) Lindquist: Op. cit., Table 13, p. 212.

15.) Burt: Op. cit., p.338.

TABLE 11. OBSERVED CORRELATIONS, THEORETICAL CORRELATIONS RESULTING FROM FIRST FACTOR SATURATIONS, AND RESIDUAL CORRELATIONS.

Tests	4	C	2	8	6	3	10	7	1	9	5	
9	.5576	-.033	-.059	-.033	.126	-.074	.015	.079	.055	-.005		<u>.263</u> -.070
5	.4719	.080	.063	-.033	-.047	.184	-.075	.011	-.090	-.094	-.070	

Below, in Table 12 , we isolate the saturation coefficients obtained in Table 10 , which were taken into account when making the final selection of tests.

TABLE 12 . GENERAL FACTOR SATURATION COEFFICIENTS.

Test No.	Coefficient	Test No.	Coefficient
C	.782	6	.741
1	.565	7	.598
2	.779	8	.756
3	.655	9	.558
4	.830	10	.612
5	.472		

The following six tests were then chosen for inclusion in the final battery:

- No. 4, Matrices,
- No. 2, Completion of Symbol Series,
- No. 8, Identical Pairs,
- No. 6, Figure Recall,
- No. 3, Classification,
- No. 1, Pintner's Four Dots.

Correlations with the criterion and saturation coefficients were clearly not the only considerations in making the final selection. So, e.g., Test No. 1 was found in practice to form an excellent test with which to establish rapport between examiner and subjects and to obtain the desired attitude and atmosphere. In this test the children are not left to themselves to carry on on their own initiative. There is continuous interaction between examiner and subjects. The presentation can be brief, the instruc-

tions are easily understood, the reaction required for every item is brief. By the nature of the test pupils are brought to the alert and to a state of concentrated attention. Most pupils enjoy the test as a kind of game. By the time it is completed, confidence has usually been established, interest awakened; the unfamiliarity of the situation and the examiner have decreased and the pupils are in a better condition to tackle the next test where they have to work on their own. For such reasons test No. 1 was included and placed first in order of presentation; and a test of similar type technically was brought in at the beginning of the second group of three tests, after the interval.

Judging by general reaction all the other tests also seemed to have high interest value and were almost always quite evidently enjoyed. In final selection, in arranging the sequence of tests and arranging the sequence of items in a test, this need for retaining sustained interest as far as possible was always kept in mind.

+ + + +

The tests retained were then subjected to further detailed study.

The percentage of correct responses on each item of each test - except for tests numbers 1 and 3, which had been well-standardized by their authors and in which no essential change had been made in the subject matter - was calculated. On the basis of the data so obtained it was possible to eliminate some items and effect some rearrangement of others. In the case of the flash-card memory test (No.6) the procedure differed somewhat. The 200 subjects were divided into six groups according to their

scores on the criterion test, and the average score of each group for each item of the memory test was determined. The results are given in Table 13.

TABLE 13. AVERAGE SCORES ON EACH OF 20 FIGURE RECALL ITEMS OBTAINED BY EACH OF SIX GROUPS.

Score on Criterion Test.	Number of Cases.	AVERAGE SCORE ON FIGURE RECALL ITEMS																			
		Item Numbers																			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
80 and more	8	2	1.25	1.63	1.38	1.63	1.63	1.50	1.63	1.50	.88	.50	1.63	1.00	.63	.88	1.63	1.25	1.00	.88	1.75
65 - 79	28	1.75	1.57	1.89	1.50	1.71	1.11	1.71	1.79	1.39	1.54	1.07	1.86	1.18	.79	1.07	1.79	.82	.71	1.11	1.07
50 - 64	43	1.84	1.44	1.67	1.37	1.67	1.07	1.67	1.67	1.14	1.23	.81	1.77	.98	.74	.88	1.44	1.19	.67	.93	.84
35 - 49	46	1.93	1.09	1.67	.70	1.33	.76	1.52	1.46	1.09	.93	.72	1.41	.96	.39	.70	1.20	.65	.39	.76	.70
20 - 34	59	1.71	1.20	1.36	.53	1.15	.71	1.59	1.34	.85	.90	.54	1.10	.92	.36	.64	.92	.59	.49	.63	.37
19 and less	16	1.38	.75	1.5	.44	1.00	.44	1.63	1.06	.63	.25	.25	.56	.81	.31	.75	.25	.38	.38	.44	.19

The same procedure was repeated with three groups instead of six. The results are in Table 14.

TABLE 14 . AVERAGE SCORES ON EACH OF 20 FIGURE RECALL ITEMS, THREE GROUPS.

Score on Criterion Test.	Number of Cases.	AVERAGE SCORE ON EACH ITEM.																			
		Item numbers																			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
65 and over	36	1.81	1.50	1.83	1.47	1.69	1.22	1.67	1.75	1.42	1.40	.94	1.81	1.14	.75	1.03	1.75	.92	.78	1.06	1.22
35 - 64	89	1.89	1.26	1.67	1.02	1.49	.91	1.60	1.56	1.11	1.08	.76	1.58	.97	.56	.79	1.31	.91	.53	.84	.76
Less than 35	75	1.64	1.11	1.39	.51	1.12	.65	1.60	1.28	.80	.76	.48	.99	.89	.35	.67	.77	.55	.47	.59	.33

These results are graphically represented in Figs. IV. 3 to IV.21, where the 6- point and 3- point graphs for a single item are superimposed, the latter being drawn in red.

In the six-group division the first group contained only 8 cases. The sharp upward bends at the top left in some of the six-point graphs were therefore not considered reliable, and disregarded. The first point in the three-point graphs clearly illustrate that little relative value should be attached to the first points of the six-point graphs.

Upward bends elsewhere, as found in the cases of Items Nos. 1, 2 and 3 were, however, taken into account, and such items discarded, as were also those showing too little general fall and hence apparently possessing too little discriminatory value, such as Nos. 7 and 13.

Items numbers 1, 2, 3, 7, 13, 15 and 17 were therefore removed. The test consequently became much shorter. In this way the danger of gradual loss of interest and concentration through fatigue or boredom was decreased, while the correlation of the test with the criterion increased somewhat.

Some of the more difficult items of tests, while not included in the final version of the test proper, were later introduced as an extra work sheet at the ends of certain of the tests with a view to later possible incorporation at the higher age levels in order to give the brighter subjects more scope. In the meanwhile they serve a useful purpose in the present form of the test in keeping these brighter pupils busy and interested while the slower ones finish the tests concerned. This is of some importance in a testing system where time limits are not imposed and is also beneficial for disciplinary considerations.

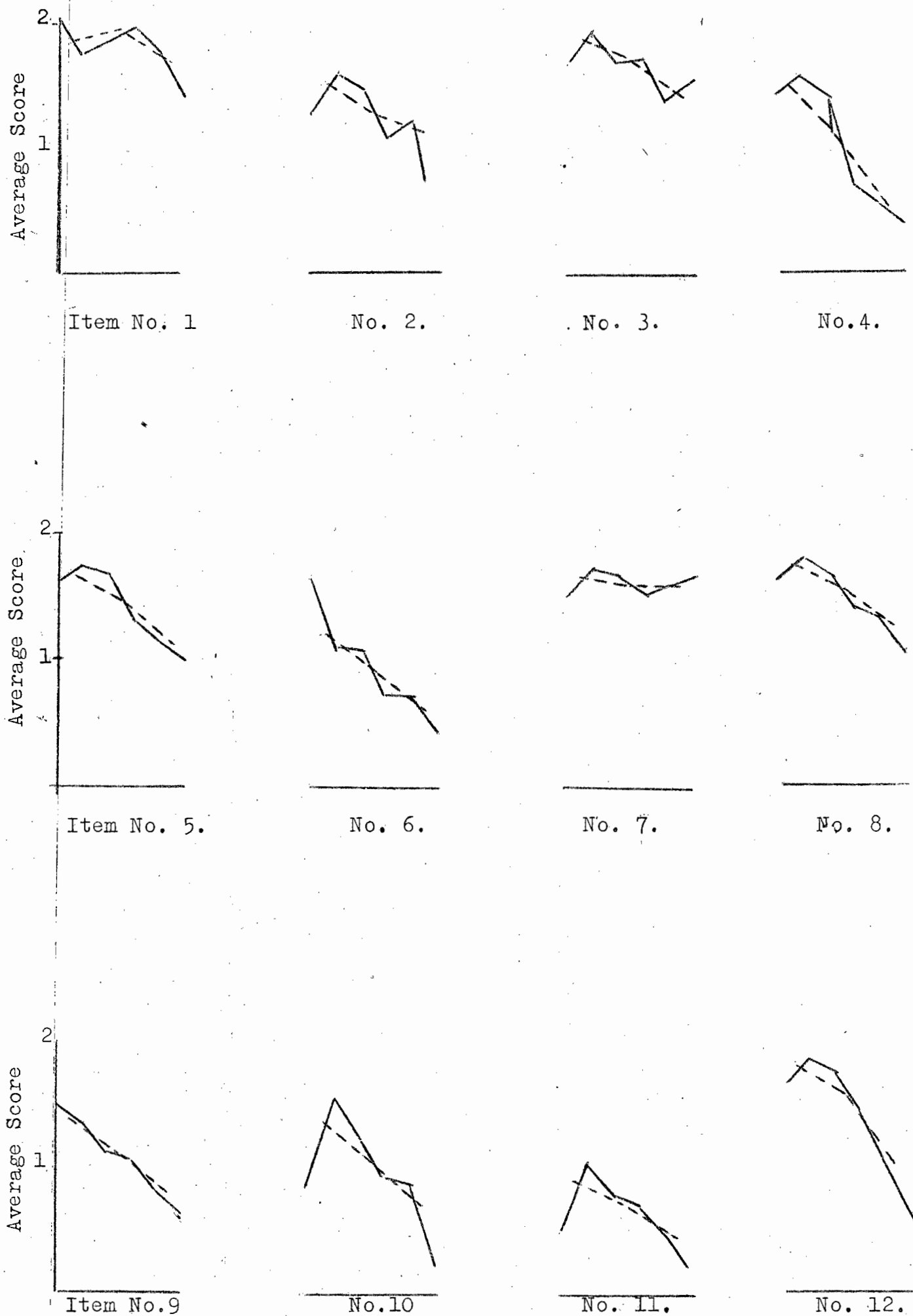


Fig. IV.3. Average Scores on Figure Recall Items, Illustrating data in Tables 13 and 14.

CHAPTER V. RELIABILITY, VALIDITY AND SOME OTHER PROBLEMS.

A coefficient of reliability was found in two ways.

First, alternate items of every subtest were allocated to two forms of the test, designated A and B. Hence two partial scores were determined for each of our 200 hearing subjects by totalling the score on forms A and B separately.

The coefficient of correlation between the A and B scores of our group was found to be

$$r_{AB} = .921 \pm .007$$

Use of the Spearman-Brown formula $\left(r = \frac{2r_{AB}}{1+r_{AB}} \right)$

yielded a coefficient of reliability given by

$$r = .959$$

It is to be noted that the splitting up of the test into the alternate forms A and B was done only for the purpose of determining the above coefficient, and that in practical use the two forms remain combined as a single test.

In the second case the reliability was tested on deaf pupils. The full test was given to a group of 60 deaf children and some three months later the same group was retested with the same test.

The coefficient of reliability so obtained, based on the repetition of the full test, is given by

$$r_{AB} = .969 \pm .003$$

+ + + + +

Different criteria were used for measuring the validity of our test, one again being used in the case of our standardizing group of hearing children, the others with

deaf children.

The South-African Group Test of Intelligence¹⁾, a test very well standardized on some 17,000 children, was the first. This, being a typical verbal test, could of course only be applied in the case of our hearing group.

The coefficient of correlation between the results of this criterion and the original raw scores on our test was found to be .741. In obtaining this coefficient, the samples drawn from different schools, as described before, were considered as a single group. Using standard scores, to be discussed in the next chapter, instead of original scores, the correlation was virtually unchanged at .742.

There is, however, another way of arriving at a coefficient of correlation in which use is made of Fisher's method of the analysis of covariance. Although we have shown before that there are no significant differences between the means of the eight sub-groups, we might as well take into account whatever differences do occur. As Lindquist²⁾ points out; "what we should compute, in situations of this kind, is the correlation within schools, rather than the total correlation".

The covariance of two variables is the mean of the products of their deviations, each from its own mean. This can be represented by

$$\frac{\sum xy}{N}$$

where x and y are the deviations of each variable from its own mean, and N the number of cases. The expression "sum of products" is used for $\sum xy$ in this sense.

Now, it can be shown that in the case of a sample consisting of a number of smaller groups, the total sum of products of

70deviations

1). Wilcocks, Op. cit.

2). Lindquist, op. cit., p. 221.

deviations from the general mean is made up of two components, one being based on the covariance within the separate groups, the other on the covariance of the group means themselves, or, differently put, one being dependent on the sum of products of deviations within a group from the individual group means, the other on the sum of products of these group means themselves—the product for each group being weighted by the number of cases contained in the group — where each mean is expressed as a deviation from the general mean. The first of these components may be called the sum of products within groups, the second, the sum of products between the groups, so that we may summarize all this by writing:

$$\begin{aligned} \text{Total Sum of products} &= \text{Sum of products within groups} \\ &+ \text{Sum of products between groups.} \end{aligned}$$

Now we want our correlation coefficient to be rid of the influence of the covariance between the means of the different schools, a covariance which is a result of the fact that our sample consisted of such different groups. Our best estimate of the true correlation between our two variables would then be based on the covariance within the different schools only (on the assumption that the correlations within these schools were homogeneous). The covariance between schools might, while included, of course have the effect of either increasing or decreasing the coefficient of correlation.

When deriving the coefficient of correlation the sum of products between schools would therefore have to be subtracted from the total sum of products to leave us with the sum of products within schools only.

The formula for the correlation coefficient

$$r = \frac{\sum xy}{N \sigma_x \sigma_y}$$

may be written as

$$r = \frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}}$$

By applying this formula to individual sub-groups and then summing for all the sub-groups, and applying the principle set out above, we have

$$r = \frac{\sum xy - \sum n \bar{x} \bar{y}}{\sqrt{\sum x^2_{\text{within groups}} \sum y^2_{\text{within groups}}}}$$

or, in words,

the coefficient of correlation

total sum of products - sum of products between groups

$$\sqrt{(\text{Sum of squares within groups for first variable}) \times (\text{ditto for the second variable})}$$

The different terms of this expression can be reduced to the following forms to facilitate calculations:

1. The total sum of products,

$$\sum xy = \sum XY - \frac{T_x T_y}{N}$$

where X and Y represent original scores, T_x and T_y the grand totals, and N the total number of cases. Again,

$$\sum XY = N \bar{x}' \bar{y}' + \sum x' y' + \bar{x}' \sum y' + \bar{y}' \sum x'$$

where \bar{x}' and \bar{y}' are the co-ordinates of arbitrary axes and x' and y' deviations from these axes.

All these quantities are taken directly from the correlation tables used earlier to obtain the coefficients given in Tables 7 and 8.

2. The sum of products between schools,

$$\sum n \bar{x} \bar{y} = \frac{T_{x_1} T_{y_1}}{n_1} - \frac{T_{x_2} T_{y_2}}{n_2} - \dots - \frac{T_{x_3} T_{y_3}}{n_3} - \frac{T_x T_y}{N}$$

where T_x is the total score of the whole first group on the criterion test, T_y the total score of the same group on our new test, etc. The last term, $\frac{T_x T_y}{N}$, which occurs here as well as in the expression for the total sum of products above, will, of course, fall away on substitution into the correlation formula, and hence may be disregarded.

The sums of squares in the denominator of the correlation formula are taken directly from the previous analysis of variance in Chapter IV.

It is to be noted that in order to be able to make use of earlier work, original raw scores were reduced in accordance with class intervals previously employed. All scores on the criterion test were therefore divided by 4, those on the new test by 5.

The totals and their products required for the between groups component are given in Table 15.

TABLE 15 . REDUCED TOTAL SCORES AND PRODUCTS REQUIRED IN THE ANALYSIS OF CO-VARIANCE (BETWEEN-SCHOOLS COMPONENT).

Schools	n(= no.of cases)	T_x (reduced total criterion score)	T_y (reduced total test score)	$\frac{T_x T_y}{n}$
1	19	211	307	3409.22
2	16	161	272	2737.00
3	20	207	344	3560.40
4	28	324	529	6121.29
5	34	361	634	6731.59
6	36	419	672	7821.33
7	30	327	544	5929.60
8	17	225	340	4500.00
Total	200			40810.53

From this table, we have therefore the sum of products between groups,

$$\sum n \bar{x} \bar{y} = \sum \frac{T_x T_y}{n} = 40810.53$$

For the total sum of products we have

$$\begin{aligned} \sum X Y &= (200 \times 12 \times 18) + 3202 + 12(42) + 18(-165). \\ &= 43936. \end{aligned}$$

And so the coefficient of correlation, after the removal of the influence of the covariance between the several school means, becomes

$$\begin{aligned} r &= \frac{43936 - 40810.53}{\sqrt{4894.20 \times 3551.64}} \\ &= .750 \end{aligned}$$

The same procedure was repeated, using standard scores instead of original raw scores of our test, yielding a still better coefficient

$$r = .773$$

To correct for attenuation we may use Spearman's formula

$$r_{\infty x \infty y} = \frac{r_{xy}}{\sqrt{r_{x, x_2} r_{y, y_2}}}$$

where r_{x, x_2} is the reliability of the criterion test, which is .93³⁾ and r_{y, y_2} the reliability of our test, which we have found to be .969.

$$\therefore r_{\infty x \infty y} = \frac{.773}{\sqrt{.93 \times .969}} = .814$$

3). Wilcocks: Op. cit.

This would be the true correlation if the two tests were both made infinitely long and therefore infinitely reliable. A coefficient so obtained would, however, not apply to the tests as they stand. A better measure of the validity of our test would therefore be to make a correction only for the degree of reliability of the criterion, so that we obtain the correlation coefficient between our test as it stands and an infinitely reliable or infallible criterion. Thus we then have a result corrected for errors in the criterion but not for errors which would appear in the new test in actual use.

In this case we would have

$$\begin{aligned}
 r_{\infty x \cdot y} &= \frac{r_{xy}}{\sqrt{r_{x, x_2}}} \\
 &= \frac{.773}{\sqrt{.93}} \\
 &= .802
 \end{aligned}$$

We now have to consider the reliability of this coefficient. When the true correlation is high, the distribution of r's from samples is skewed and the use of the probable error method itself not reliable⁴⁾.

According to Fisher, the function

$$z = \frac{1}{2} \log_e \frac{1+r}{1-r}$$

is almost normally distributed, however, for any value of r. The standard deviation of the distribution of z is given by

$$\sigma_z = \frac{1}{\sqrt{N-3}}$$

In this case (r = .802), using Lindquist's Table 14' (p.215) we find

$$Z = 1.104228$$

75...../We

4). Lindquist, op. cit. p.212

We next have to determine the limiting values of r between which we could be confident that no less than 99% of such r 's will lie, or within which we could be confident "at the one percent level" that the true r lies.

Of the area under the normal curve 99 percent lies within 2.576 σ above and below the mean. This means that the difference between the observed and "limiting" z , expressed as a normal deviate, must be 2.576,

i.e., $Z_H - Z_o = 2.576 \sigma_z$ (where Z_H denotes the hypothetical, limiting value of z , and Z_o the observed value.)

and, as

$$\sigma_z = \frac{1}{\sqrt{200 - 3}}$$
$$= .0712 ,$$

$$\therefore Z_H = 1.1042 \pm (2.576 \times .0712)$$
$$= 1.2876 \text{ or } .9208$$

The values of r corresponding to these values of Z , obtained from the same table, are .858 and .726

Hence, $r = .802 + .056$ }
 or $-.076$ } at the one percent level.

+ + + + +

For our second test of validity use was made of teachers' judgment of the intelligence of their pupils. This test was applied to the results obtained from deaf children.

Each class teacher was asked to classify his pupils in five categories, viz., very bright, bright, average,

dull, very weak. These pupils were also classified in five similar categories on the basis of the test results, and the coefficient of contingency for these two systems of classification determined from Pearson's formula⁵⁾:

$$C = \sqrt{\frac{\phi^2}{1 + \phi^2}}$$

$$= \sqrt{\frac{\chi^2}{N + \chi^2}}$$

where

$$\phi = \frac{\chi^2}{N} = \frac{1}{N} \sum \left[\frac{\left(f_{xy} - \frac{f_x f_y}{N} \right)^2}{\frac{f_x f_y}{N}} \right]$$

This formula is reduced by Holzinger⁶⁾ to the form

$$C = \sqrt{\frac{S-1}{S}}$$

where

$$S = \sum \left\{ \frac{f_{xy}^2}{f_x f_y} \right\}$$

This was done for each school for the deaf separately, as the standards of classification differed so widely between schools that it was out of the question to pool the results.

The coefficients of contingency obtained are given in Table 16.

77/Table 16.

- 5). Pearson, K : "On the Theory of Contingency and its Relation to Association and Normal Correlation", Draper's Research Memoirs, Biometric Series I, 1904.
- 6). Holzinger, K: "Statistical Methods for Students in Education", Ginn & Co., 1928.

TABLE 16 . COEFFICIENTS OF MEAN SQUARE CONTINGENCY BETWEEN TEST RESULTS AND TEACHERS' JUDGMENT, IN THE CASE OF DEAF CHILDREN.

School	Number of pupils tested:	Number of Classes:	Race:	Coefficient of Contingency:
St Vincent School Johannesburg	87	10	European	.603
Skool vir Dowe's Worcester	192	22	European	.490
School for the Coloured Deaf, Worcester	90	8	Non-European (various)	.636
Dominican School for the Deaf, Wittebome.	174	18	Non-European (various)	.514
Kutlwanong School for the Deaf, Roodepoort.	28	3	Bantu	.690

The probable error was worked out in one case, the first named in the Table, as an example, and found to be .0066⁷⁾

There are several points to be noted about these findings. Some of these are the following:

1. The maximum value that the coefficient of contingency can attain.

- 7). The formula for the probable error of the contingency used was as follows:

$$\frac{.6745}{\sqrt{N}} \left[\frac{\psi^3 + 1 - \phi^2}{(1 + \phi^2)^3} \right]^{\frac{1}{2}}$$

where
$$\phi^2 = \frac{1}{N} \sum \left[\frac{(f_{xy} - f_x f_y / N)^2}{f_x f_y / N} \right] = S - 1$$

and
$$\psi^3 = \frac{1}{N} \sum \left[\frac{(f_{xy} - f_x f_y / N)^3}{(f_x f_y / N)^2} \right]$$

attain in a five-by-five table is .894.

2. In the case of any one school we are dealing with the subjective impressions of different teachers of different classes. Hence there is no common standard of judgment. Standards differ as a result of individual differences in intelligence, temperament, character, experience of teachers, as well as of differences in the average ability in different classes. Obviously the effect of this diversity of judging standards will increase with an increase in the number of classes and hence of judges concerned in the classification of the different groups. This is very evident in the case of the second school in the above table when compared with the others.

3. There is a further particular complication arising in a school for the deaf. In these schools the main and primary goal of all teaching can be regarded as language attainment: to give the pupil the ability to read, to lipread, to speak, to express himself in language symbols, whether written or spoken. As all teaching is thus concentrated on and around language, it is consciously or subconsciously by language attainments that the child is primarily judged by his teacher. Now, it is the social and sociable type of deaf pupil who picks up language much more readily than his more reticent fellow. It is this sociable type, extraverted, full of self-confidence, imitative by nature, eager to please, unhampered by self-conscious shyness and fear of blundering, who almost invariably makes a better impression than his native intelligence would warrant. Conversely, the reticent, shy, self-conscious child is very often intellectually superior but realises his defect more, is more difficult to reach and to draw out, thus gets less practice and makes in general a poorer showing in speech and other language attainments.

This fact is pointed out and emphasized by Snyders-Oomen⁸⁾
This was also our own experience when discussing individual cases with teachers and enquiring into the grounds for the respective classifications. So, for example, we found in one instance, where fifteen cases showing a wide divergence between teachers' and test classifications were discussed with the teachers concerned, six cases where good or poor speech was cited as the reason for the classification. On further enquiry it was often found that a child classified as of poor intelligence on account of poor speech was doing very well in arithmetic and vice versa! (In passing it may be mentioned that amongst the remaining nine of the above fifteen pupils, there were cases of poor eyesight, "difficult" behaviour and other irrelevant factors which were evidently influencing the judgment on intelligence).

For reasons similar to the above, the child with some useful residual hearing is very often placed in a higher category than he deserves as a result of better language development.

4. As many deaf children first come to school much later than normal, there is often a wide range of ages in the same class; the tendency is definitely for the range to be wider than in schools for normal children. While the classification according to test results was based on I.Q.'s derived from these results⁹⁾, so that the age factor was eliminated, many teachers seemed unable to make due allowance for these age differences in forming their estimates of intelligence. We had some flagrant examples of this. If

8) Snyders-Oomen, Op. cit., p.68.

9) See next Chapter.

A cannot reach the same standard of achievement as B in the same class, who may be five years older, A is judged to be inferior to B in intelligence although he may be doing very much better than B relatively to his age. It is interesting to compare Terman's findings in this connection. He found the correlation between I.Q.'s on the Stanford Revision of the Binet-Simon tests and teachers' rankings on a five-point scale to be .48, about which he remarks: "This is about what others have found, and is both high enough and low enough to be significant"¹⁰⁾ On further investigation he found that the serious mistakes in cases of disagreement between test results and teachers' ratings were nearly all on the part of teachers where children either over or under age for their school class were concerned. "In estimating children's intelligence... the teachers often failed to take account of the age factor", he states, and concludes that "the disagreements between the tests and the teachers' estimates are thus found, when analyzed, to confirm the validity of the test method rather than to bring it under suspicion".

5. It was also found, as mentioned before in passing, that the child causing disciplinary trouble was placed lower than he deserved for self-evident even if involuntary and subconscious reasons.

Other factors influencing such a subjective criterion as teachers' estimates might be mentioned but the above are perhaps some of the most important.

Taking all this into account, the contingencies found were deemed to be highly satisfactory as confirmation of the validity of our test.

81/As

10) Terman, L.M. The Measurement of Intelligence, Harrap, London 1937, p.75, 76.

As a third test of validity the results of the group test were compared with results from the Snyders-Oomen individual test, for a group of 60 deaf pupils.

This yielded a coefficient of correlation

$$r = .746 \pm .050$$

(The correlation between the South-African Individual Scale and the South-African Group test is very similar:

$$r = .7398 \pm .0365^{11}) .)$$

+++++

We also wished to determine to what extent that which we were measuring was the product of education and training instead of the native capacity we were trying to probe. This problem is of especial significance in case of the deaf, with whom the degree of command of verbal language is largely bound up with the duration of school attendance.

Those who have only just come have no notion of verbal language and therefore do not possess the language symbols which are often regarded as indispensable for the development of abstract thought. If mental development is retarded by lack of verbal language, we would expect a decrease in this retardation as the child gains command of language and practice in its use. This would mean a tendency to improvement in the I.Q.'s with increase in the duration of the educational process.

Coefficients of correlation between I.Q.'s and the period the child has been attending school were determined with a view to throwing some light on this problem, and are presented in Table 17.

82/Table 17

11). Fick, M.L: An Individual scale of General Intelligence for South-Africa, Bureau of Educational Research Pretoria.

TABLE 17. CORRELATION BETWEEN I.Q.'s AND DURATION OF SCHOOL ATTENDANCE. *)

School	Number of Cases	Coefficient of Correlation.	Value of r required for significance	
			At the 5 percent level	At the 1 percent level
St Vincent Johannesburg.	80	-.026	.220	.287
Skool vir Dowes Worcester	180	.0603	.148	.193

Only those who have been at school not less than one year but not more than twelve years have been included.

The correlation coefficients obtained are quite evidently completely insignificant.

From these results we conclude that that which our test measures is not to any appreciable extent influenced by schooling or even by language mastery.

The problem of the influence of language was also studied from a different angle.

Pupils in a school for the deaf may always be classified into two categories: those who acquired any language they possess entirely through artificial teaching and training and those who acquired their language by natural means. The first class includes all those born completely deaf for all practical purposes, those who lost their hearing and so their language very early, usually before their sixth year, and those with some residual hearing which, however, they do not utilise for some reason or other - probably poor intelligence. The second class

includes those hard of hearing pupils whose residual hearing enabled them to acquire at least some language, however restricted and defective, in a normal way. Amongst these may be found children with a greater loss of hearing than some included in the first category, but utilising their little hearing to a greater extent, possibly, one would surmise, as a result of a more alert intelligence. (The presence of this type here would seem to favour a somewhat higher average I.Q. for this category.) Furthermore, we have here all pupils in hard of hearing classes receiving normal instruction through hearing aids. This second class also includes children who, although completely deaf, became so after having acquired their language normally, and retained it notwithstanding possible deterioration in speech and absence of further development. This is usually the case with those who became deaf after their sixth year.

The difference between these two categories of the deaf is profound and is the result of all that language as a natural instrument of thought, communication and expression implies.

Biserial r between I.Q.'s and these two categories of pupils was calculated for one hundred pupils in the Worcester School for the Deaf, which was the only school available with a fair number of pupils with natural language. Of the group of 100 there were 38 possessing naturally acquired language, while the remaining 62 had acquired their language entirely by artificial means. We found $r_{bis} = .1601$ in favour of the group with artificial language. For significance at the 5 percent and 1 percent levels the values of r would have had to be at least .197 and .256 respectively. If any positive relationship actually did exist, one would have expected it to be in favour of the group with natural

language, as pointed out above. The result obtained, while superficially seeming to point in the opposite direction, merely serves to indicate the absence of any relationship.

Again we seem justified in concluding that our test appears to be entirely independent of language ability in so far as such ability is dependent on the sense of hearing.

The distribution of scores of the standardizing group of 200 hearing children was skewed as shown in Table 18 and Fig. VI. 1.

TABLE 18 . DISTRIBUTION OF RAW SCORES OF 200 HEARING SUBJECTS.

S c o r e	Number	S c o r e	Number.
Below 28	1	78 - 87	26
28 - 37	3	88 - 97	36
38 - 47	2	98 - 107	38
48 - 57	12	108 - 117	35
58 - 67	17	118 - 127	12
68 - 77	16	128 and over	<u>2</u>
		Total	200

This was a result of the inclusion of a fairly large proportion of less difficult items, which was done on purpose to give even the weaker deaf children a chance of gaining confidence by fostering a sense of achievement at the outset.

This skewing also caused an apparent tendency for the regression between these results and those on the criterion

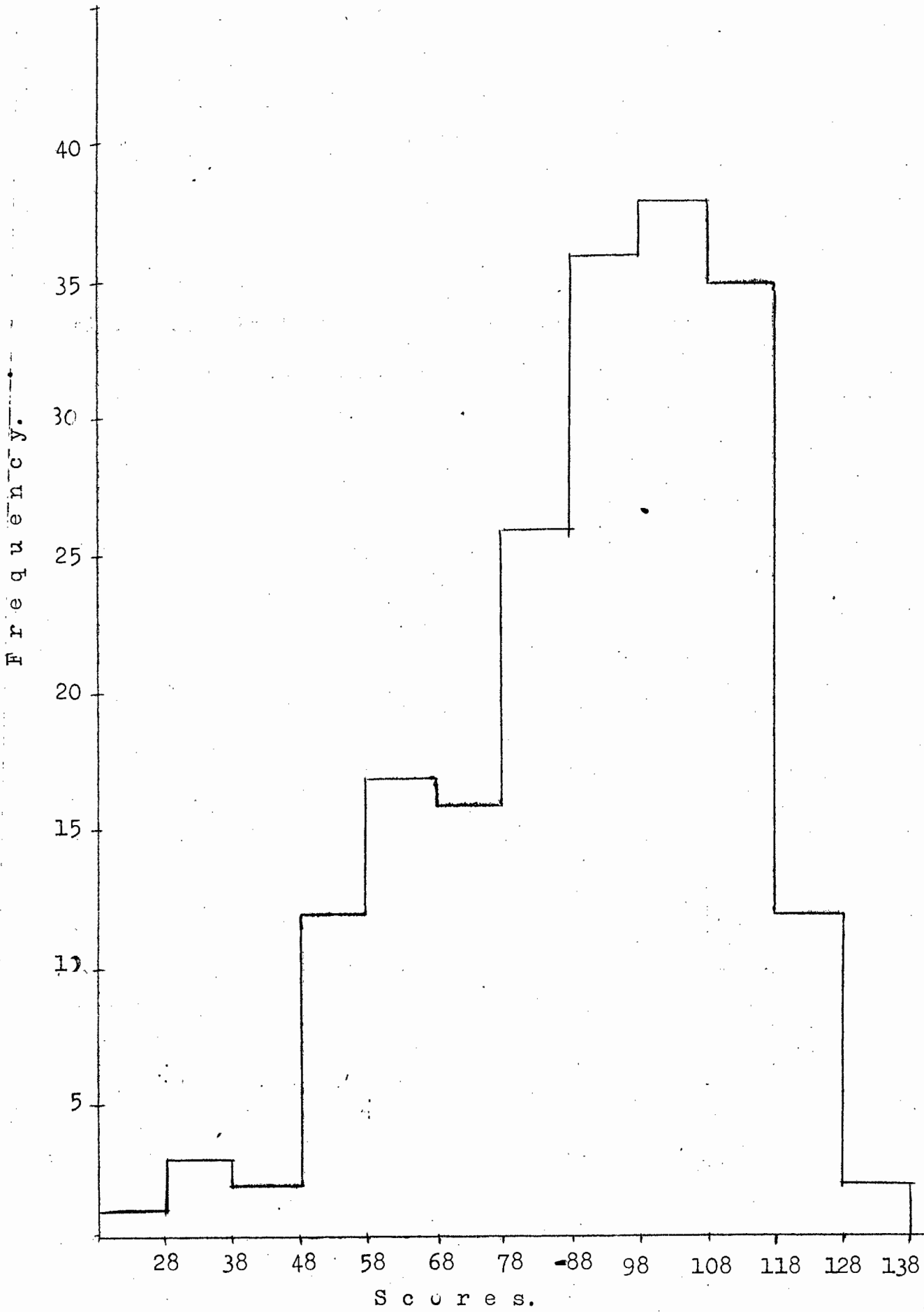


Fig. VI.1. Distribution of Raw Scores of 200 Hearing children.

test to be somewhat curvilinear ¹⁾ as a result of some congestion in the higher score region. To correct this skewing, even if slight, raw scores were transmuted into standard T scores by McCall's method ²⁾, assuming a normal distribution of the actual variable being tested, using the following procedure:

From the percentages of cases obtaining the different total scores, the areas under the normal curve and hence the normal deviates corresponding to such scores were determined; to this was added 5 in order to eliminate negative signs, and the results were multiplied by 10.

The standard scores so obtained (see Table 19) were plotted against the original scores and the resulting graph smoothed somewhat where necessary (Fig. VI.2). Smoothed standard scores were read off from this graph, rounded and tabulated in Table 20.

88/Table 19

- 1). For our original data a correlation ratio, $\eta = .794$ was obtained as compared with the correlation coefficient, $r = .741$. These values applied in the short version of Blakeman's formula yielded a value of 3.684 for $\frac{4.047}{\sqrt{N}} \sqrt{\eta^2 - r^2}$, which is less than 4.047, the critical value. Again, applied in the full formula, we found the left hand side $\eta^2 - r^2 = .0679$, while the right hand side $\frac{4.047}{\sqrt{N}} \sqrt{(\eta^2 - r^2) \{ (1 - \eta^2)^2 - (1 - r^2)^2 + 1 \}}$ reduces to .07252, which is more than the value of the l.h.s. We would therefore be justified to regard the regression, even in the case of the original data, as reasonably rectilinear. (Cf. Holzinger, op. cit., p.183).
- 2). McCall: How to Measure in Education, McMillan, New York, 1922.

TABLE 19 .STANDARD SCORES (BEFORE SMOOTHING) CORRESPONDING TO ORIGINAL RAW SCORES.

Original Score	Stand. Score	Original Score	Stand. Score	Original Score	Stand. Score	Original Score	Stand. Score
23	24	54	35	85	46	116	63
24	24	55	35	86	46	117	64
25	24	56	36	87	47	118	65
26	24	57	36	88	47	119	65
27	24	58	37	89	48	120	65
28	24	59	37	90	48	121	66
29	27	60	37	91	49	122	66
30	27	61	37	92	49	123	67
31	27	62	37	93	49	124	67
32	27	63	37	94	50	125	68
33	27	64	39	95	50	126	70
34	27	65	39	96	51	127	70
35	28	66	39	97	51	128	73
36	28	67	40	98	52	129	73
37	28	68	41	99	52	130	76
38	29	69	41	100	52	131	76
39	29	70	41	101	53	132	76
40	29	71	41	102	53		
41	29	72	41	103	53		
42	30	73	41	104	54		
43	30	74	42	105	55		
44	31	75	42	106	55		
45	31	76	43	107	56		
46	31	77	43	108	57		
47	31	78	43	109	57		
48	31	79	43	110	58		
49	32	80	44	111	59		
50	32	81	44	112	59		
51	32	82	44	113	60		
52	33	83	44	114	60		
53	34	84	46	115	61		

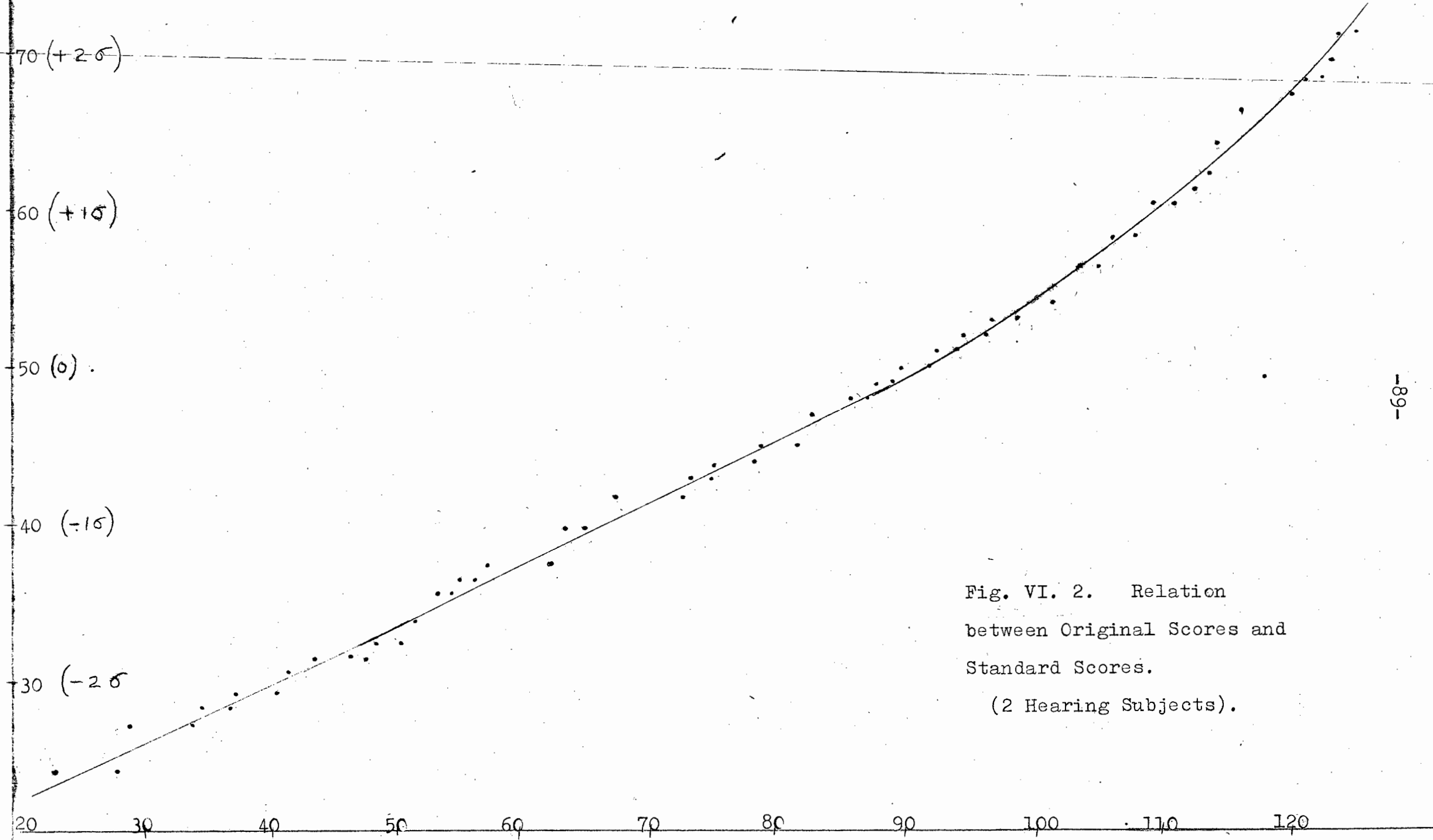


Fig. VI. 2. Relation
 between Original Scores and
 Standard Scores.
 (2 Hearing Subjects).

TABLE 20 . SMOOTHED STANDARD SCORES CORRESPONDING TO ORIGINAL
RAW SCORES.

Original Score	Stand. Score	Original Score	Stand. Score	Original Score	Stand. Score	Original Score	Stand. Score	Original Score	Stand. Score
15	21	45	31	75	42	105	55	135	78
16	21	46	31	76	42	106	55	136	79
17	21	47	32	77	43	107	56	137	80
18	22	48	32	78	43	108	56	138	81
19	22	49	32	79	43	109	57	139	82
20	22	50	33	80	44	110	58	140	83
21	23	51	33	81	44	111	58	141	84
22	23	52	33	82	44	112	59	142	85
23	24	53	34	83	45	113	60	143	86
24	24	54	34	84	45	114	60	144	87
25	24	55	35	85	46	115	61	145	88
26	25	56	35	86	46	116	62	146	89
27	25	57	35	87	47	117	63	147	90
28	25	58	36	88	47	118	63	148	91
29	26	59	36	89	48	119	64	149	92
30	26	60	36	90	48	120	65	150	93
31	26	61	37	91	48	121	66		
32	27	62	37	92	49	122	66		
33	27	63	37	93	49	123	67		
34	27	64	38	94	50	124	68		
35	28	65	38	95	50	125	69		
36	28	66	39	96	50	126	70		
37	28	67	39	97	51	127	71		
38	29	68	39	98	51	128	72		
39	29	69	40	99	52	129	73		
40	29	70	40	100	52	130	74		
41	30	71	40	101	53	131	74		
42	30	72	41	102	53	132	75		
43	30	73	41	103	54	133	76		
44	31	74	41	104	54	134	77		

The distribution of the standard scores of our 200 hearing subjects is given in Table 21. and in Fig. VI.3, which shows the skewing to be much less than before, while the scatter diagram appeared on inspection to conform better than before to a straight regression line. The correlation with the criterion test, by the method of covariance, increased slightly, from .750 to .773.

TABLE 21. DISTRIBUTION OF THE STANDARD SCORES OF 200 HEARING SUBJECTS.

S c o r e :	N u m b e r :	S c o r e :	N u m b e r :
Below 24	1	48 - 53	46
24 - 29	3	54 - 59	41
30 - 35	14	60 - 65	20
36 - 41	22	66 - 71	10
42 - 47	41	72 - 77	<u>2</u>
		Total	200

The correlation between standard and original raw scores was found to be represented by a coefficient of .985.

The scores on each subtest was also scaled in a similar way and the standard scores so obtained for the different subtests then totalled to obtain another version of a standard general total for each child. The results so obtained were found to agree very closely with those obtained before. The correlation with the criterion test remained virtually unchanged and the rectilinearity of the regression was not at all improved. This alternative method was discarded on account of the complication and increased labour it introduced into scoring in practice, hence the voluminous figures are not included here. We found that standard scores as obtained from the full totals gave the relative standing of

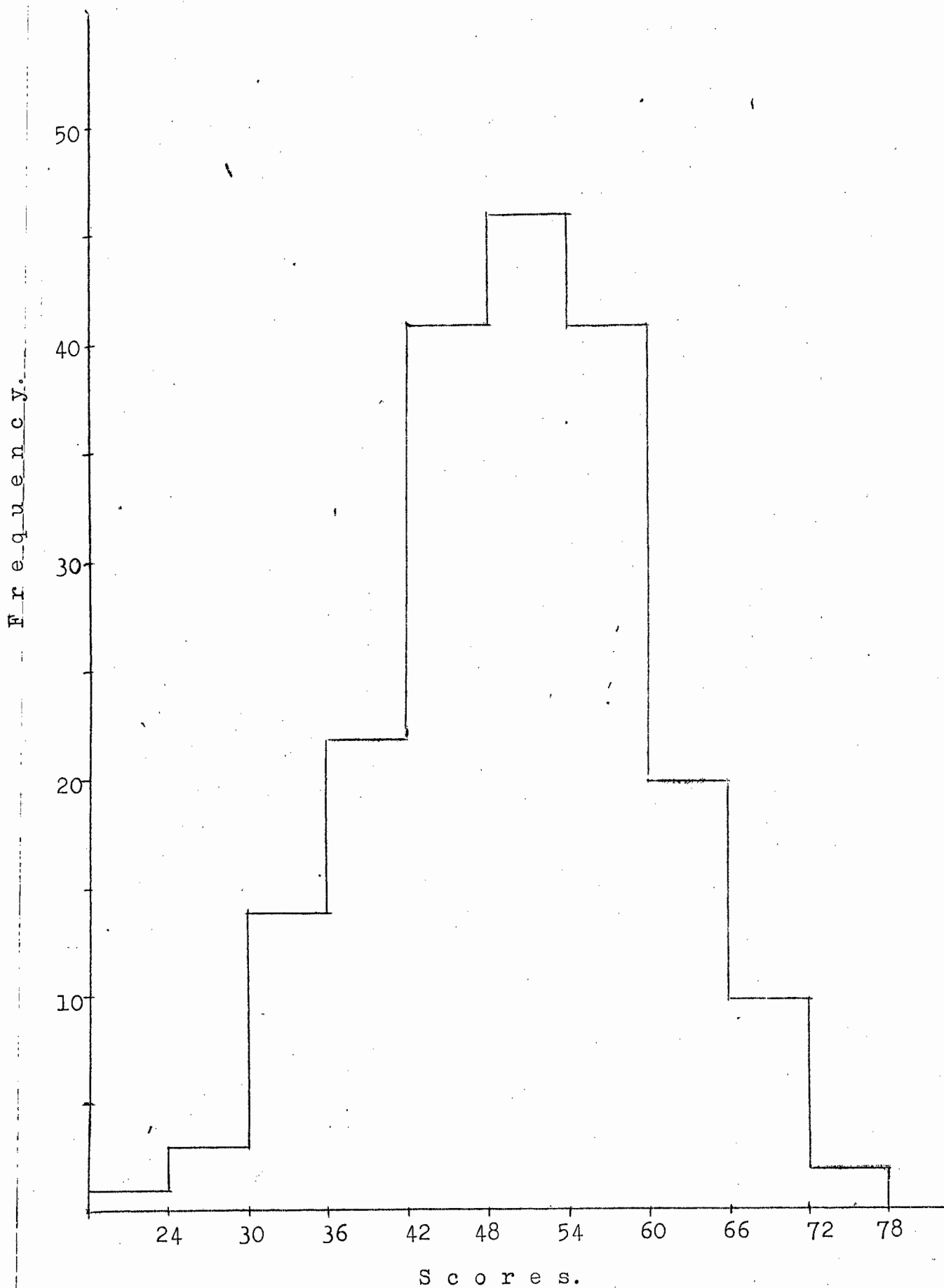


Fig. VI. 3. Distribution of Standard Scores of 200 Hearing Children.

93...../pupils

pupils with sufficient accuracy for all practical purposes.

+ + + + +

The most useful way in which to express the results of a general intelligence test is still in the form of Intelligence Quotients, provided the information conveyed by such a single index is regarded, with the necessary reservations, for what it is worth, and not misused to draw unwarranted conclusions. Unfortunately - for our present purpose - there are not enough deaf pupils available in South-African schools for the deaf for the empirical standardization of norms for the different ages.

To obtain provisional norms of I.Q.'s for our test two different methods were employed.

The first method was based on the regression of the new test standard scores on the criterion scores as found with our standardizing group of hearing subjects. We have already shown (footnote, page 86) that the data even in original form may be regarded as yielding linear regression. This is even more pronounced when using standard scores.

The coefficient of regression could be obtained in the usual way by substituting in the formula.

$$b_{xy} = r_{xy} \frac{\sigma_x}{\sigma_y}$$

the values of the right-hand factors. (These values were obtained in the course of calculations in connection with the correlational work earlier.)

We therefore have

$$b_{xy} = .773 \frac{20.04}{9.738}$$

$$= 1.591$$

Our best estimate of the true regression, since we are justified in assuming our eight sub-groups to be random samples from the same population (see Chapter IV), would however be the average regression within groups, expressed in the form:

$$b_{xy} = \frac{\sum \text{products within groups}}{\sum \text{squares for the } Y\text{-distribution, within groups.}}$$

$$= \frac{\sum \text{products (total)} - \sum \text{products between groups}}{\sum \text{squares}_y \text{ within groups}}$$

Substituting the values, based on standard scores and full criterion scores, that were used for arriving at the value $r = .773$ in Chapter V, we have

$$b_{xy} = \frac{29349.08}{18390.96}$$

$$= 1.596$$

The regression equation then becomes

$$X = 1.596y + a.$$

Substituting the values of the means, through which the line passes³⁾, we have

$$44.500 = (1.596 \times 49.615) + a .$$

whence $a = - 34.7,$

so that the equation finally reduces to

$$X = 1.596 y - 34.7.$$

3). G.U. Yule: An Introduction to the Theory of Statistics, 13th edition, London, 1948, p. 313.

The graph of this equation is given in Fig. VI.4.

By means of this equation, the scores of our test can be transmuted into scores on the criterion test, for which a full set of I.Q. norms is available for all month-age-groups from 120 to 192 months.

Our regression line was based on the chronological age groups 10 to 12 years only, but was assumed to hold for other ages as well (see discussion in connection with the second method.)

This method is to a much greater extent dependent on the criterion test than the second method, to be described further on. One disadvantage was that in this case we could not proceed below the ten-year limit as the criterion norms only go down to that age group. This was one of the reasons why this method was discarded in favour of the second method. Consequently the full set of norms based on the regression-method is not presented here. However, the coefficient of correlation between I.Q.'s obtained by this method and those obtained by means of the later method, for 100 cases ranging in age from 120 to 192 months, was found to be .962, although in extreme cases, i.e. at high I.Q.'s and for low ages there was some noticeable divergence, differences of up to eleven points between results from the two methods being found.

In the case of the second method, deviations, in terms of normal deviates from the medians of the respective age-groups, were expressed in a form similar to the conventional I.Q.

The procedure was as follows:

Sub-groups were formed of the chronological age groups of 10, 11, and 12 year-old children of our standardizing group. (Chronological ages of two cases were not available, so that

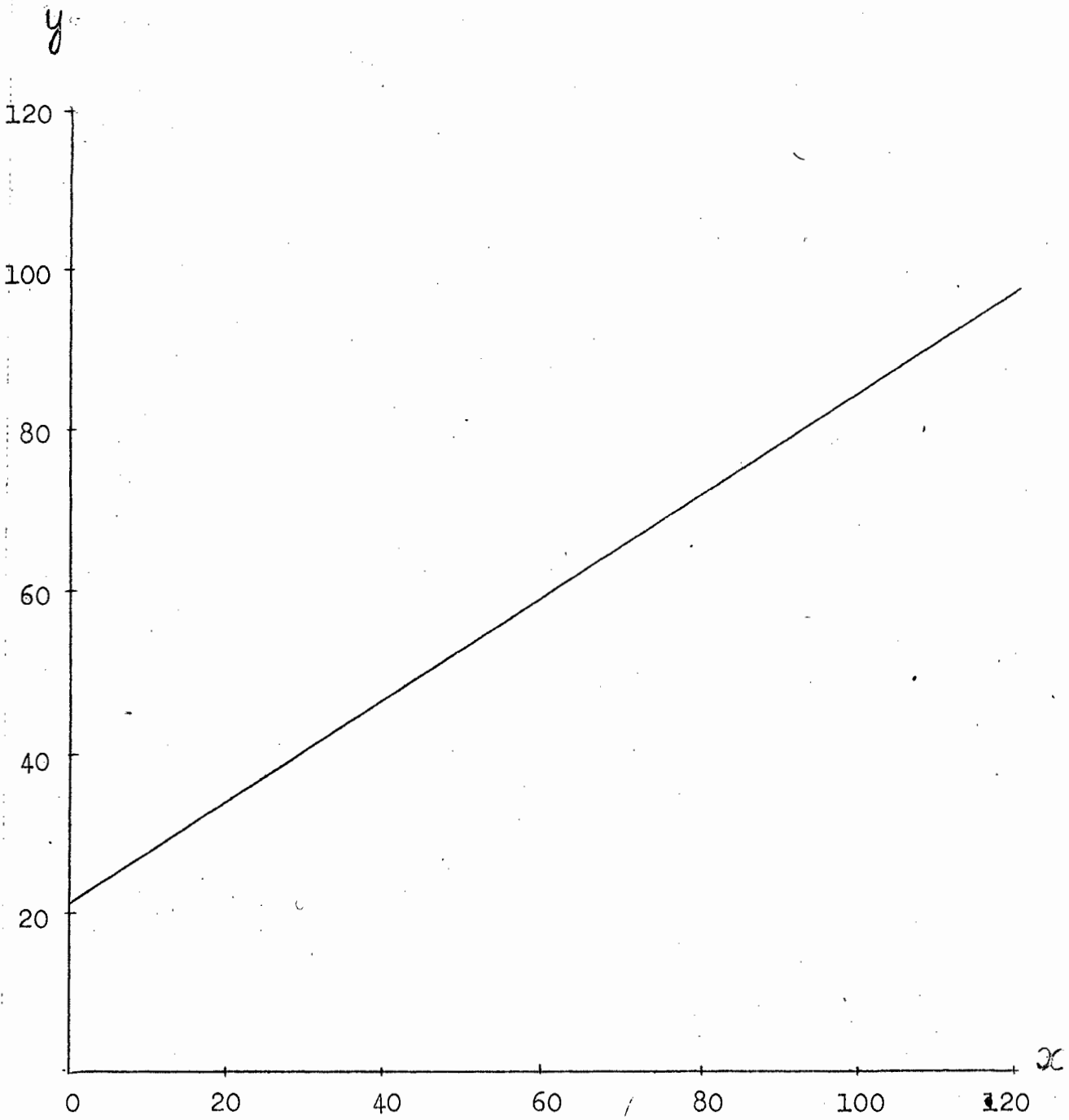


Fig. VI. 4. Graph of Regression Equation:

$$X = 1.596 y - 34.7$$

the total number used was 198). The median mental age of each of these groups was determined by means of their scores on the criterion test, for which age norms are available. The median standard scores on our new test were then considered to be the score representative of each of these three mental ages. Such a score for such an age would therefore represent an I.Q. of 100. The data are given in Table 22.

TABLE 22. MEDIAN MENTAL AGES AND SCORES FOR SUB-GROUPS.

Sub-group.	No. of Cases.	Median Mental Age.	Median Standard Score.
1	59	132 months	45
2	75	155 "	51.5
3	64	164 "	54

A graph was then constructed with ages in months measured along the abscissa and standard scores, together with the corresponding original raw scores along the ordinate. The three points representing the median scores for the three median mental ages of our sub-groups were plotted and found to lie in an almost perfectly straight line. This line was named the I.Q. 100-line.

The frequency of the occurrence of different I.Q.'s in an unselected sample of children has been determined by Terman⁴⁾. The normal deviates from the median corresponding to these frequencies expressed as percentages were read from the well-known tables of the normal curve. In this

98...../way

4). Terman, L.M: The Measurement of Intelligence, London, 1937, p. 78.

way it was possible to determine the normal deviates corresponding to the different I.Q.'s. E.g., an I.Q. of 92 was found to correspond to a deviation of $-.675\sigma$. Such a deviation was then marked off below the median points of the three sub-groups, the straight line joining them being named the I.Q.-92 line. By simple interpolation the standard scores representing an I.Q. of 92 for the different intermediate ages lying between the three known ages could be read off from the scale on the ordinate. These values could be transmuted into original raw scores by means of Table 20. However, the original score values were now also marked off on the ordinate of the graph, so that the values could be read off directly. By a similar procedure other lines representing other I.Q. values were drawn.

Purely tentatively these lines were produced beyond the limits of the median mental ages of our three groups, and, on the assumption that they would remain straight, extrapolated values could be found. The graph obtained in this way and from which I.Q.'s corresponding to different scores at different ages can be read off, is presented in Fig. VI.5 in simplified form. The complete form will be found in Part II with the test.

This graph, and values derived from it, are offered here as provisional approximations until such time as empirical norms can be established. Nevertheless, I.Q.'s found by means of the graph for deaf pupils of all ages between approximately 8 and 16 years were used for the determination of the coefficients of contingency presented before, in Chapter V, and which proved satisfactory.

It should be noted that "I.Q.'s" obtained in this way, by comparing the score of the individual with the performances of others of the same age instead of with the average performances of other age groups, do not quite conform

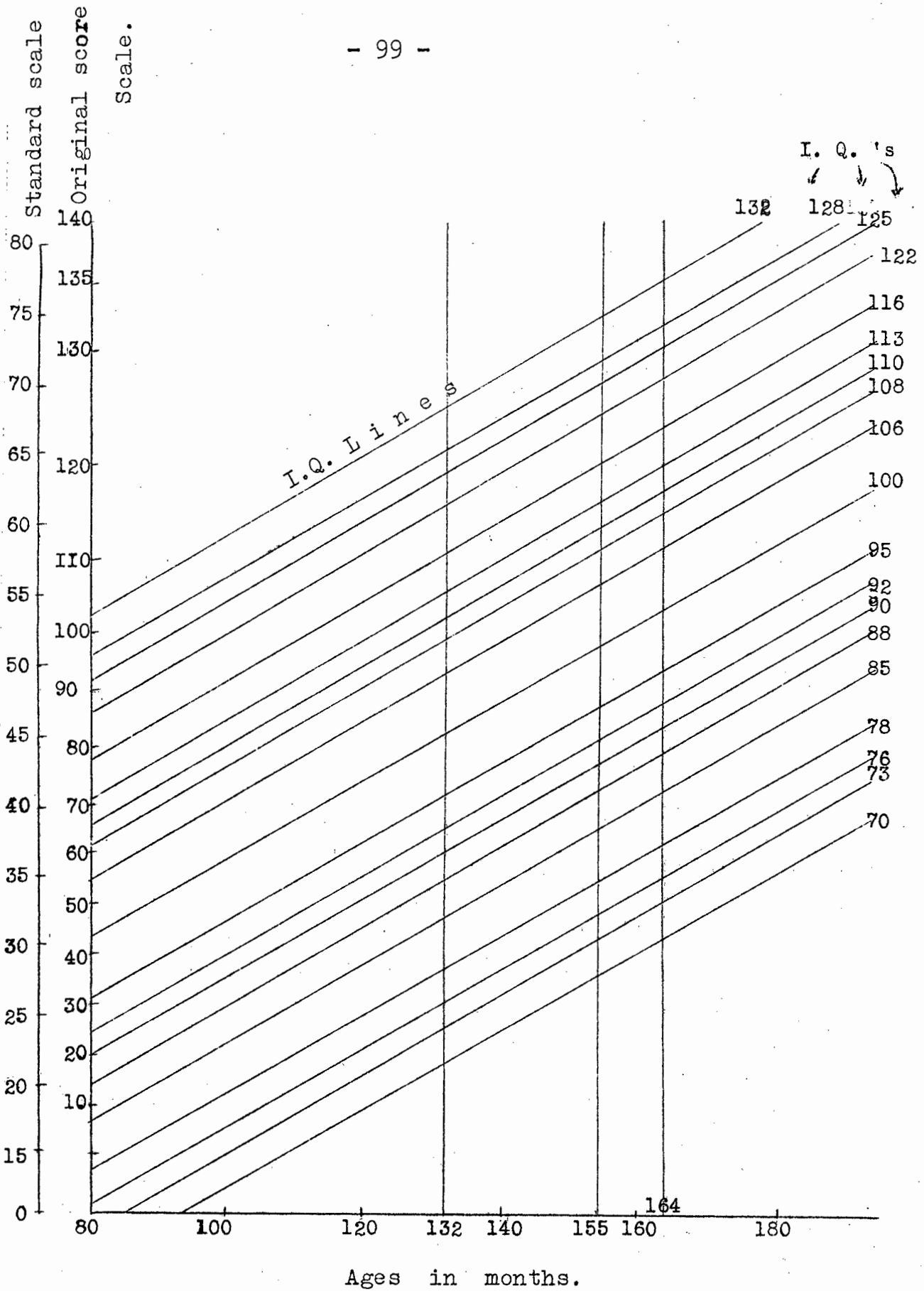


Fig. VI. Scores Corresponding to Different I.Q.'s at Different Ages.

100...../to

to the definition of the intelligence quotient. For example, a child of 140 months, scoring 110, would according to these norms, have an "I.Q." of 114. His mental age would however be 177 months, so that, from the definition, his I.Q. should be $\frac{177}{140} \times 100$, i.e., 126. This discrepancy seems to be a consequence of accepting Terman's table of the frequency of occurrence of the different I.Q.'s. For this reason we were careful to state that our results are expressed in a form similar or analogous to the conventional I.Q., not identical with it.

Something similar occurs in the norms for the South-African Group Test, which were derived in a somewhat similar manner, apparently also making use of Terman's findings. Taking as a random example the case of a child of 120 months with a weighted total score of 62, and referring to the graph of norms⁵⁾, we find an I.Q. of 130. From the same graph it appears, however, that 62 is the average score for 192 months, so that this child has a mental age of 192 months, according to which the I.Q. should be $\frac{192}{120} \times 100$, i.e. 160, by definition.

It is not our intention to discuss this problem in full at this point, merely to point out the fact of its existence, and to have the nature of our norms clearly understood.

There are certain advantages attached to this method. If the youngest group tested in standardizing is, as in our case, the 10 year age group, it would be impossible, using the ordinary I.Q. formula, to assign an I.Q. to a 10 year old child scoring below the average, as there would be no

101...../norms

5). Wilcocks, : The South African Group Test of Intelligence, Description and Directions, (Second Revised Edition), Pro Ecclesia, Stellenbosch, 1942.

norms of average I.Q.'s below that age by which to determine his mental age. Similarly it would not be possible to determine I.Q.'s of children in the highest standardizing age group who score above the average of their group. These limitations do not occur in the present method.

The distribution of the I.Q.'s - obtained by this method - of 289 European deaf children between the ages of 8 and 16 years inclusively, is given in Table 23. and Fig. VI.6, where a normal curve is superimposed. The mean of this group of deaf pupils was found to be 98.533 - an interesting finding in the light of the controversy about the intelligence level of the deaf as discussed in Chapter I - , and the standard deviation 14.52. The application of the test of goodness of fit yielded the value 11.3108 for χ^2 , from which, since 8 frequency

TABLE 23 . DISTRIBUTION OF I.Q.'s OF 289 DEAF PUPILS, 8 TO 16 YEARS OF AGE (DIFFERENT SCHOOLS TOGETHER).

I.Q.	Number of Cases.	Percentage of Total.	I.Q.	Number of Cases	Percentage of Total
Below 70	9	3.1	100-109	84	29.1
70-79	22	7.6	110-119	39	13.5
80-89	32	11.1	120-129	20	6.9
90-99	80	27.7	130 and over	3	1.0
			Total	289	100

groups were used, and by the use of Pearson's Table XII⁶⁾

102...../Fig. VI. 6.

6). As reproduced by Holzinger, op.cit., Table 59, p.248.

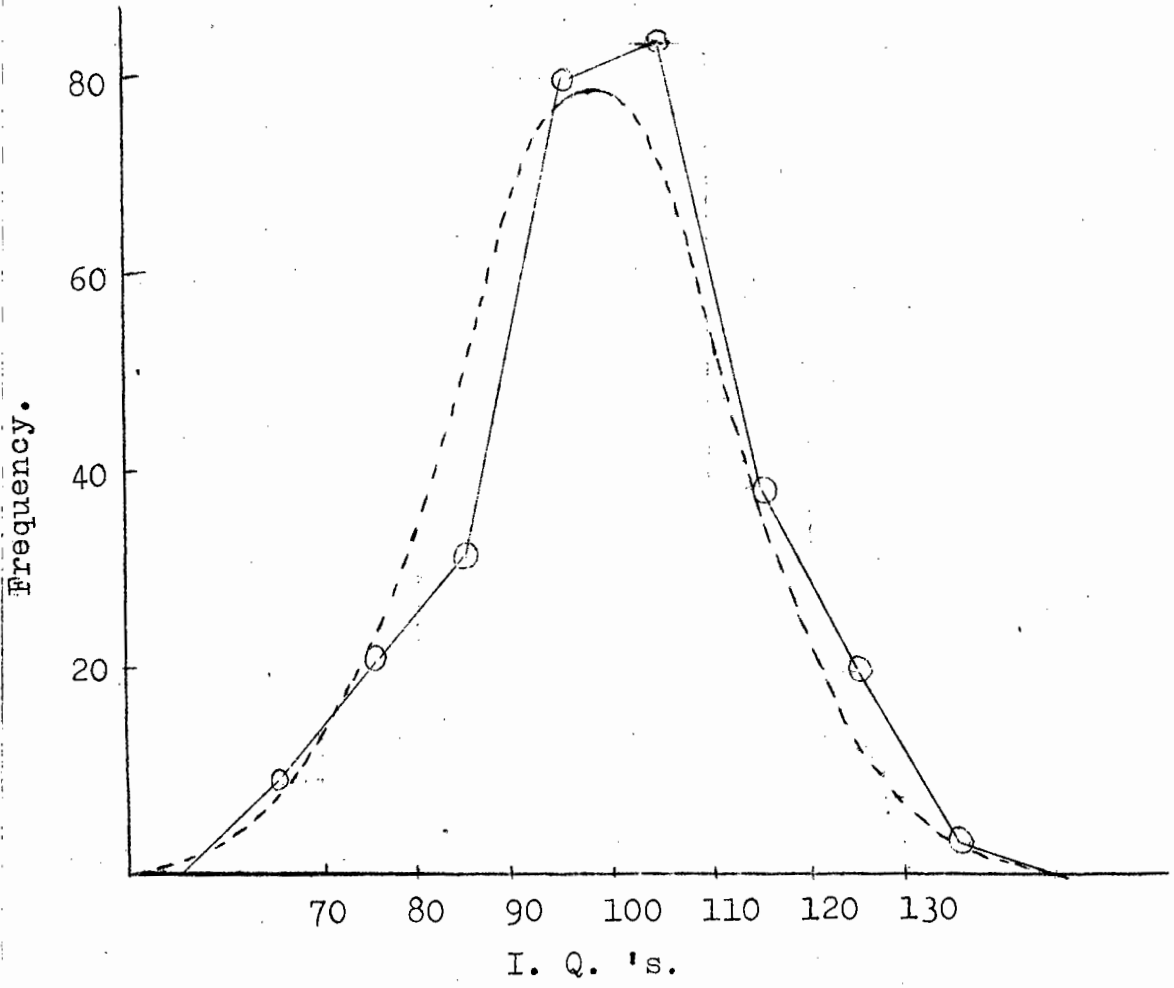


Fig. VI. 6. Distribution of I.Q.'s of 289 Deaf Children (European).
(Mean: 98.533, S.D.; 14.52)

we find the value of $P = .128$. This means, in Holzinger's words, that in 128 cases out of 1,000 trials we should get, in random sampling, a fit as bad or worse than that which would be obtained if the real distribution were represented by the normal curve we fitted. While this fit cannot be regarded as good, since the value of P is less than .2, it is certainly not so bad as to warrant complete rejection of the applicability of the normal curve to the data.

Pearson's method of moments⁷⁾ was then applied to a distribution of 307 scores of deaf children (the same group as above with the addition of 18 cases of another school). The first four unadjusted moments were found to be

$$\begin{aligned} \nu_1 &= 0 \\ \nu_2 &= 1.926 \\ \nu_3 &= .5395 \\ \nu_4 &= 11.7809. \end{aligned}$$

After adjusting by Sheppard's corrections,⁸⁾ the moments became

$$\begin{aligned} \mu_1 &= 0 \\ \mu_2 &= 1.8793 \\ \mu_3 &= .5395 \\ \mu_4 &= 10.8288. \end{aligned}$$

From these Pearson's Constants were derived and found to be:

$$\beta_1 = .04385 \left. \begin{array}{l} \text{and } \beta_2 = 3.066 \pm .036. \\ \pm .018 \end{array} \right\}$$

104...../Taking

7). Holzinger, op. cit., p.338 f.

8). Sheppard, W.F., Proceedings of the London Mathematical Society, Vol.XXIX, as quoted by Holzinger.

Taking as criteria for a normal curve⁹⁾ that

$$\beta_1 \text{ should be } = 0$$
$$\text{and } \beta_2 \text{ should be } = 3$$

we find that our values, while very nearly satisfying the second condition, do not comply sufficiently well with the first.

A measure for kurtosis is given by

$$\gamma = \beta_2 - 3$$

which here becomes $3.066 - 3 = .066$, an insignificant value, so that we may conclude that in this respect our distribution is quite normal.

The value of β_1 found indicates some appreciable skewing (the expression $\frac{\sqrt{\beta_1} (\beta_2 + 3)}{2 (5\beta_2 - 6\beta_1 - 9)}$ may be taken as a measure of skewness, its value here being $.1047$)¹⁰⁾

A possible explanation for this skewing is to be found in the fact that our distribution consisted of the scores of pupils of three different schools, between which there might conceivably be some systematic difference.

We therefore studied the case of the largest sub-group, consisting of pupils of a single school, by itself. The

105/distribution

9). Brown & Thomson: The Essentials of Mental Measurement, Cambridge, 1921.

10). Calculation of Pearson's further constants, k_1 and k_2 yielded the values .00045 and .02184, from which it appeared that a curve of Type IV would provide a better fit. This does not differ very much from the normal curve.

distribution of scores for this school is given in Table 24 and Fig. VI. 7. (Pupils in special classes for the mentally retarded were excluded).

TABLE 24 .DISTRIBUTION OF SCORES OF 180 PUPILS FROM ONE SCHOOL.

I.Q.	No of Cases.	Percentage.	I.Q.	No of Cases.	Percentage.
Below 70	3	1.67	100-109	57	31.67
70-79	9	5	110-119	22	12.22
80-89	20	11.11	120-129	10	5.55
90-99	57	31.67	130 and over	2	1.11
			Total	180	100

In this case the mean was 99.96, the standard deviation 12.98 and the value of χ^2 6.954. The value of P corresponding to this χ^2 and eight frequency groups was found in Pearson's Table of P to be .436. According to Holzinger¹¹⁾ a P of .2 or more can be taken as indicating a good fit. We may therefore conclude that in this case the distribution of I.Q.'s, as obtained by means of our norms, very satisfactorily conforms to a normal distribution. This provides further vindication of our method of derivation of norms.

In Table 25 and Fig. VI.8 the actual average scores obtained by the different age-groups of the European deaf children tested are compared with the average (I.Q. 100) scores for the same age groups as read from our graph of norms. 106/ Fig. VI. 7.

11). Holzinger, op. cit., p.247

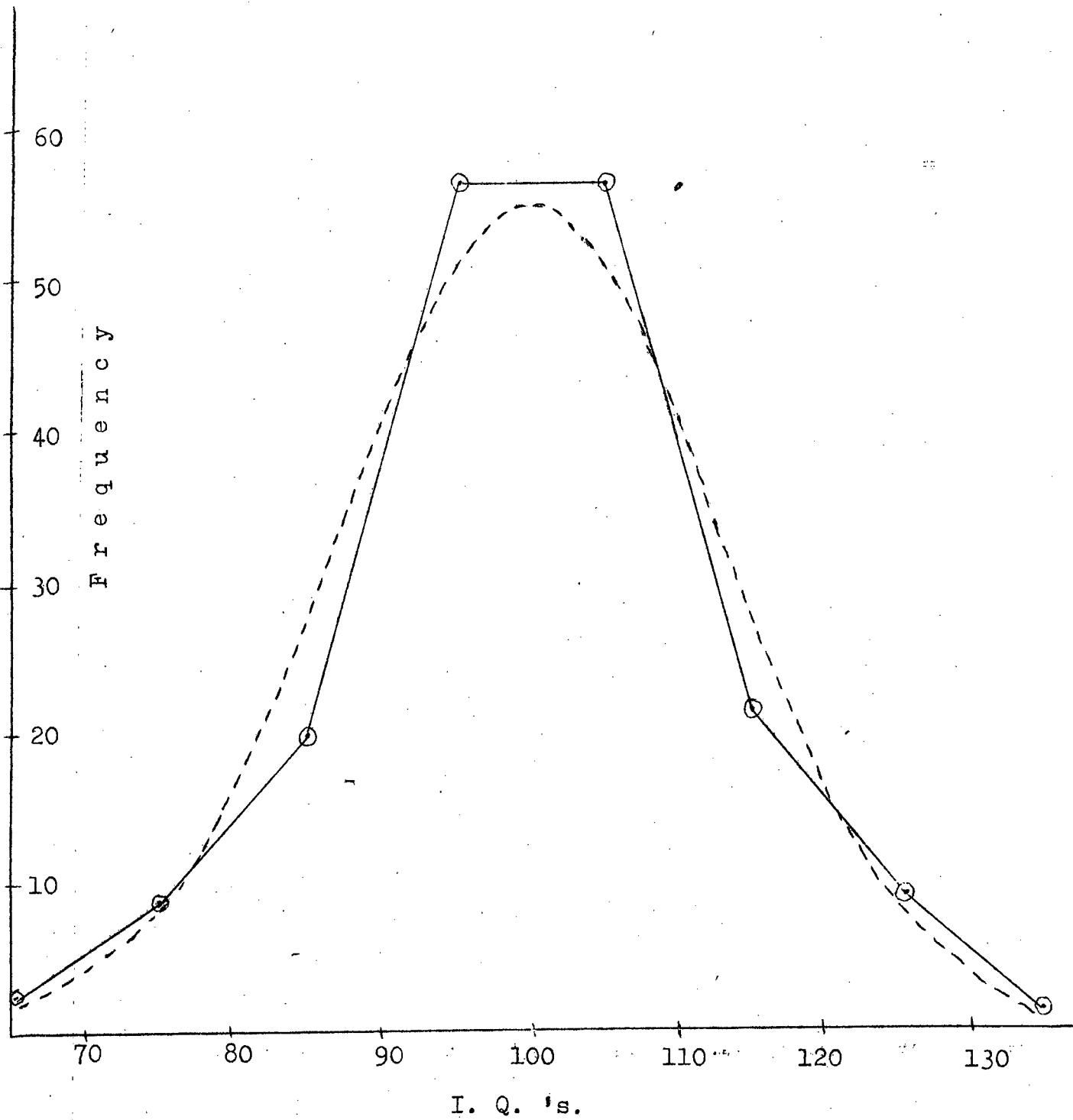


Fig. VI. 7. Distribution of I.Q.'s. of 180 Deaf Pupils of One School (European).

(Mean: 99.96, S.D.: 12.98).

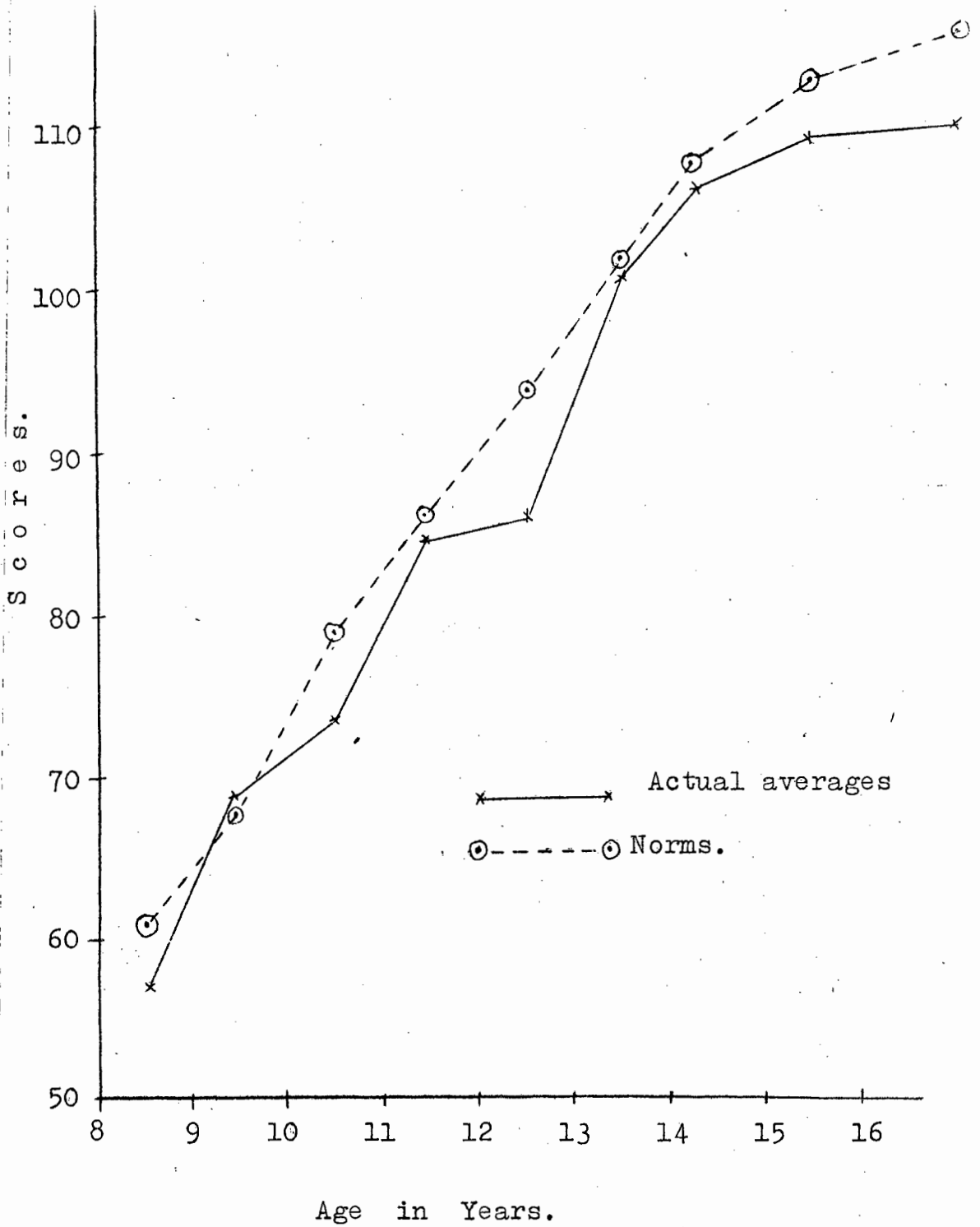


Fig. VI. 8. Actual Average Scores obtained by Deaf Pupils, compared with norms.

TABLE 25. MEAN SCORES OF DIFFERENT AGE GROUPS OF EUROPEAN DEAF CHILDREN COMPARED WITH NORMS FOR THOSE AGES.

Age Groups.	Average Age in months	Number of Cases.	Actual mean Scores obtained	Mean Score as in Norm-Graph.	Difference between last two columns	Standard Deviation of actual scores
8 yrs.	102.8	22	57.4	61	-3.6	4.715
9 yrs.	113.3	19	69	68	+1	6.359
10 yrs.	126	25	73.6	79	-5.4	6.025
11 yrs.	137.6	21	85	86.5	-1.5	4.967
12 yrs.	150.2	24	86.5	94	-7.5	3.545
13 yrs.	162	32	100.8	102	-1.2	4.521
14 yrs.	173.7	29	106.2	108	-1.8	3.817
15 yrs.	185.8	26	109.2	113	-3.8	3.206
16 yrs. and over	-	78	111.7	116	-4.3	2.172

Only in one case, the 12 year group, the difference between the mean of the scores actually obtained and the norm for the same age-group exceeds twice the standard deviation of the actual scores of the group. The general trend of agreement is as close as can be expected with the small numbers in the different groups, and considering the fact that we are comparing deaf groups with norms deduced from hearing subjects. It is to be noted that the differences between norms obtained by extrapolation (i.e., by producing the I.Q. lines beyond the limits of our standardizing groups) and the corresponding actual averages are no greater than the differences in the case of the 10, 11 and 12 year groups, which were the chronological age groups of our original standardizing sample of hearing children, and for which ages the test was originally intended, nor than the

differences in the case of 132 months, 155 months and 164 months, the median mental ages of our standardizing sub-groups.

We conclude that the provisional norms for all ages, from eight to sixteen years, render reasonably useful results in practice.

It should be clearly understood that the norms as presented here are based on those of the South-African Group Test of Intelligence, which, at the time this research was under way, was the best standardized test available for South African children.

Should any adjustment in the norms of the latter test eventually prove necessary, the norms of our test would be influenced and would have to be adjusted accordingly. In the meanwhile our norms should be regarded in the light of the deliberate acceptance of the South African Group Test norms, as they stand at present.

PART II ---- THE TEST.

P A R T II.

I N T R O D U C T I O N .

It is perhaps necessary once again to urge great caution in the interpretation of test results.

Even with a test of reliability as high as that of the present one, any one at all conversant with psychological testing should know that in any individual case a quite considerable variation in the quality of performance on different occasions may occur, no matter what measures are taken to standardize circumstances and procedure.

It can never be emphasized enough that no binding decisions, affecting an individual's whole course of life, such as exclusion from school on grounds of mental deficiency, or placing in a special class for the mentally retarded, or even labelling a child as dull - or bright, for that matter, - should be based on the results of a single application of a particular test. Yet this is a course of action that is still too often followed in our schools and appears to be routine procedure even with some psychologists.

It is not fair to the child.

It is not scientifically correct.

It is simply not good enough.

We cannot here discuss the impossible and unfortunate situations that sometimes arise in consequence, nor the irreparable harm the child may suffer as a result.

The psychologist, even the author of a test, should always retain a wholesome critical, even somewhat sceptical,

attitude towards all tests, including his own. Intelligence tests are useful instruments, but they are not infallible. The test score is the outcome of the interaction in a given situation of many factors, of which the test itself and the hypothetical "native" intelligence of the subject are two, but by no means the only two.

Wrong decisions arrived at on the basis of an incorrect test result may not be very frequent. When they do occur we may dispose of the matter as an unfortunate exception, with a shrug and a "sorry, just too bad" - attitude. The individual child concerned, however, has only his single life to be made or marred. And this life cannot be repeated after having been spoiled by an inexcusable mistake on the part of authorities that should have known better.

A group test of intelligence is primarily intended as a survey measure, to give a general indication of the intellectual capacities of the group and its distribution in the group. It has never claimed the ability to give a final verdict in individual cases, and should not be so used. The group test is, however, very useful in showing up those cases that need further investigation.

We recommend the following as routine procedure in testing, whether of deaf or of hearing school children:

1. A group test, such as the present one, is applied to the class or group under standard conditions.
2. The class teacher's own judgement, based on impressions derived from daily contact and observation of the children as well as on written class examinations is also recorded.

3. All extreme cases, i.e. either borderline or below normal (or exceptionally brilliant, if desired) are set aside for further tests.

4. All those cases that show a marked discrepancy between teachers' ratings and test results, especially where a good subject, according to test results, is doing poor work, according to teachers' ratings, are also marked for further study.

5. The special cases under 3. and 4. are now all given individual intelligence tests, preferably at least two different tests at an interval of about one month. Conditions of health, possible environmental circumstances engendering emotional disturbance and tension, and any other factors that may adversely influence achievement on tests, are in the meantime considered, and, where possible, alleviated.

Any final conclusions about the intellectual status with implications for placement in classes, changes in curriculum, decisions about future courses of study or choice of vocation in these special cases should always be based on the results of no fewer than three different tests taken at fairly long intervals, together with the considered opinion of teachers concerned. Afterwards such cases should be re-examined at least annually, always with an open mind and willingness to reconsider the previous verdict.

Finally, let the obvious be stated once more: that success and progress in any individual case is not a function of intelligence only; that there are many other factors and aspects of the personality that need to be taken into consideration; and that any decision about a child's future should take account of the whole picture.

NAME.....

DATE TO-DAY.....

CLASS.....

DATE OF BIRTH.....

SCHOOL.....

1.

1.



2.



3.



4.



5.



6.



7.



8.



9.



10.



11.

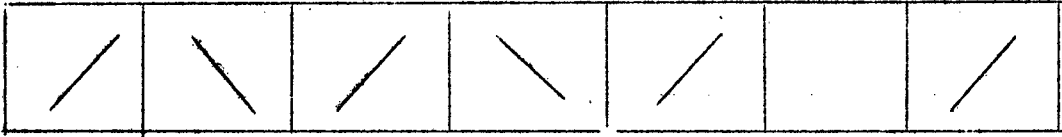
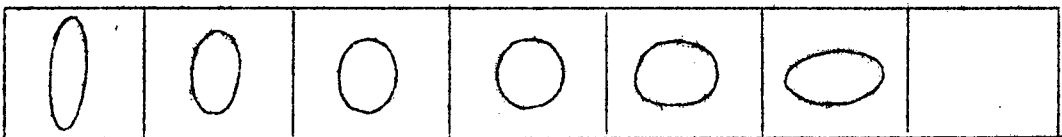
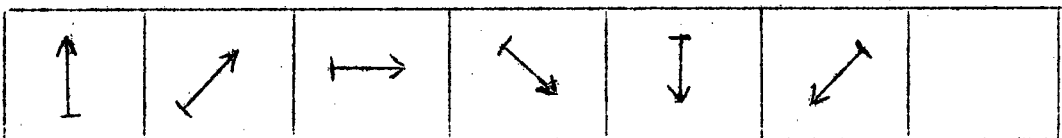
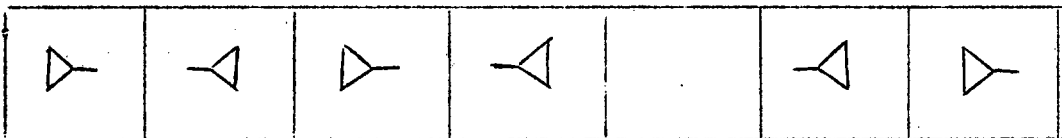
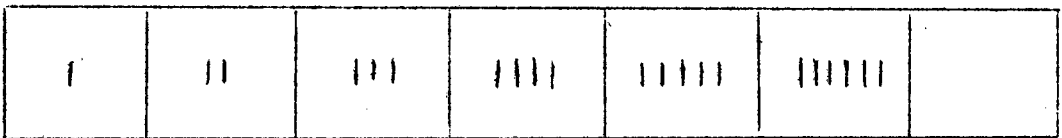
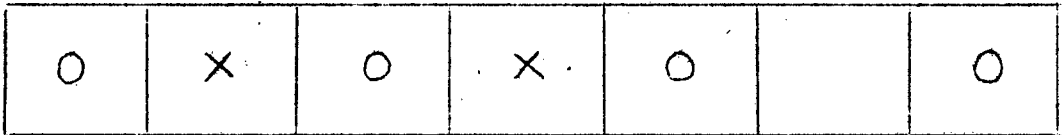
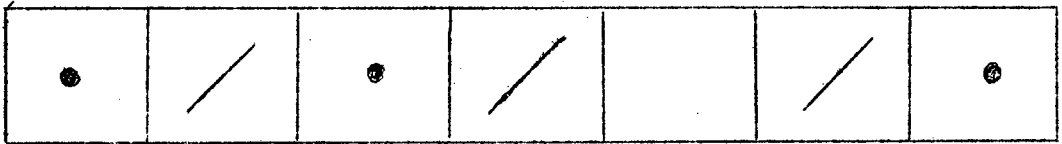


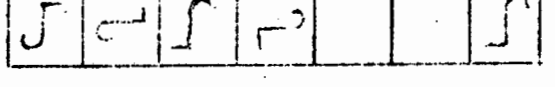
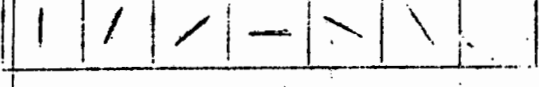
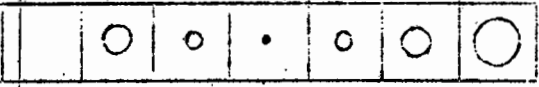
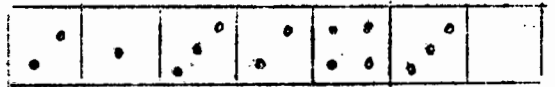
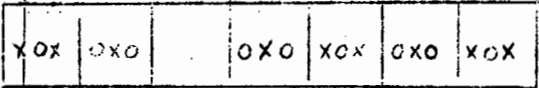
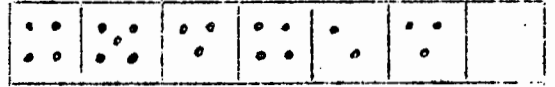
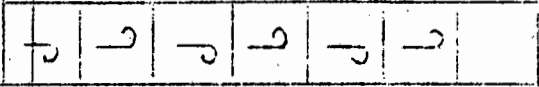
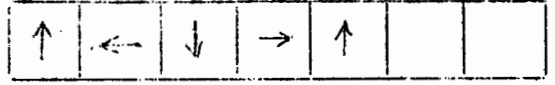
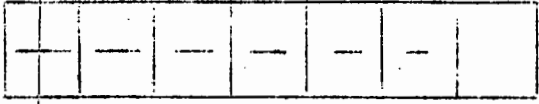
12.



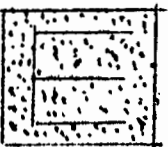
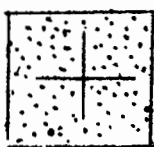
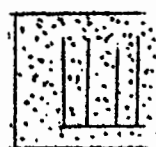
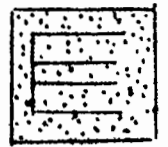
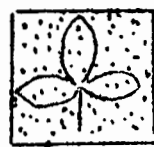
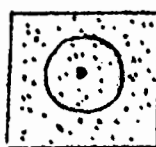
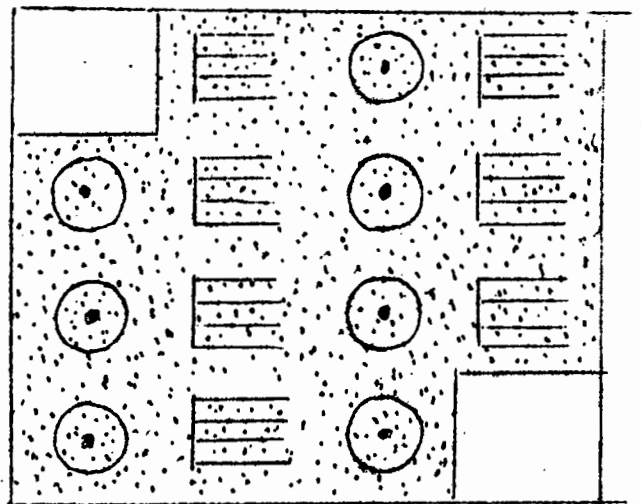
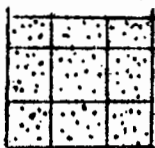
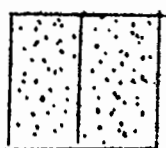
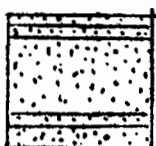
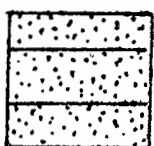
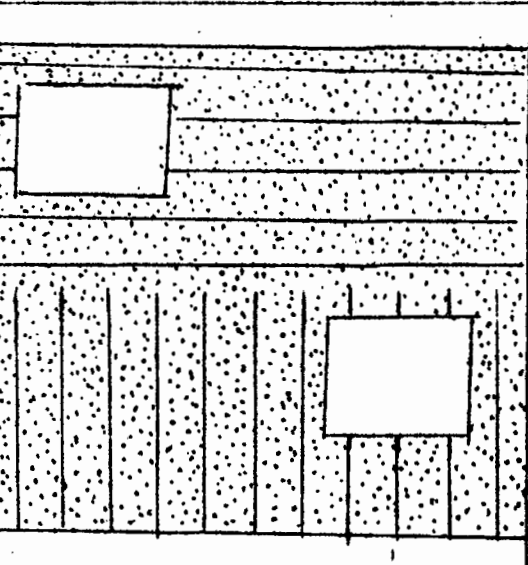
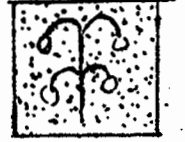
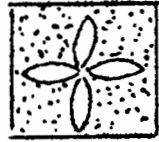
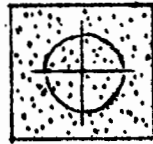
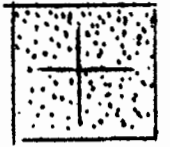
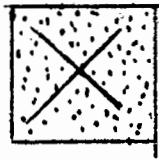
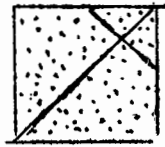
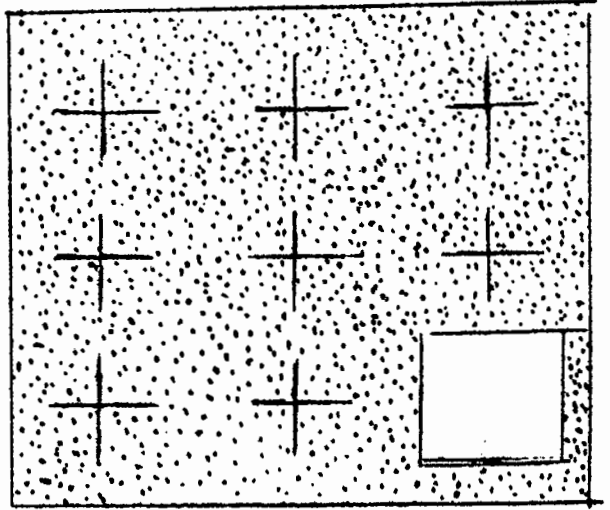
2.

PRACTICE PAGE.

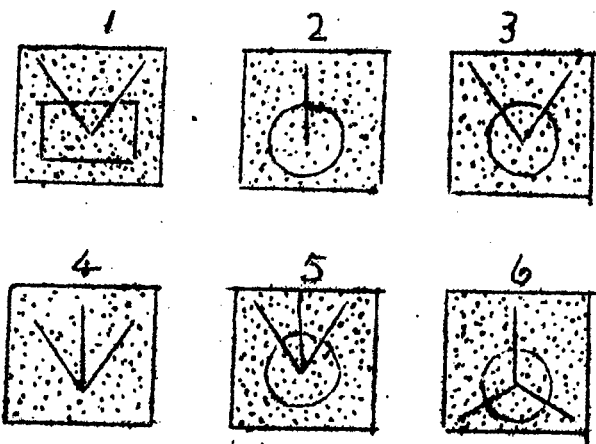
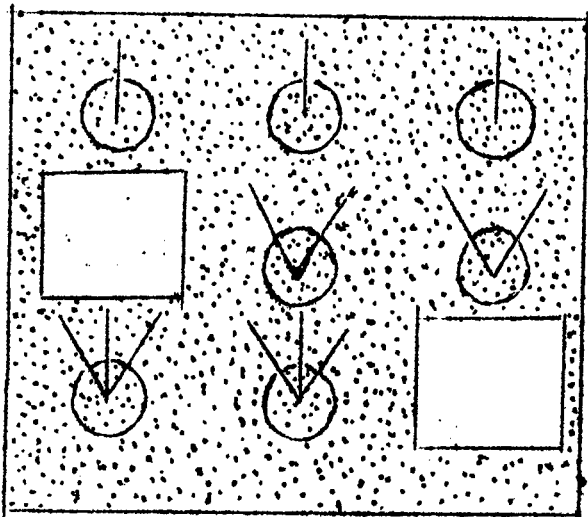




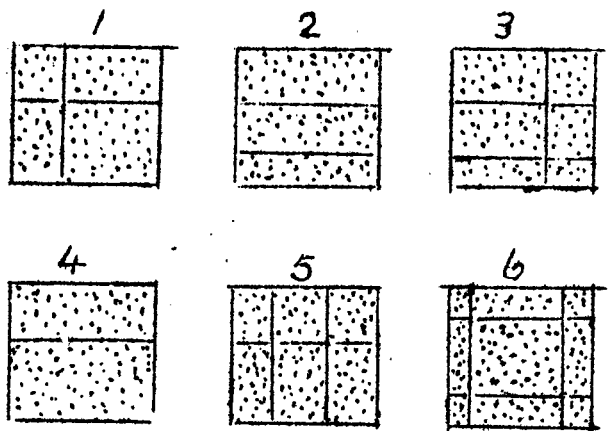
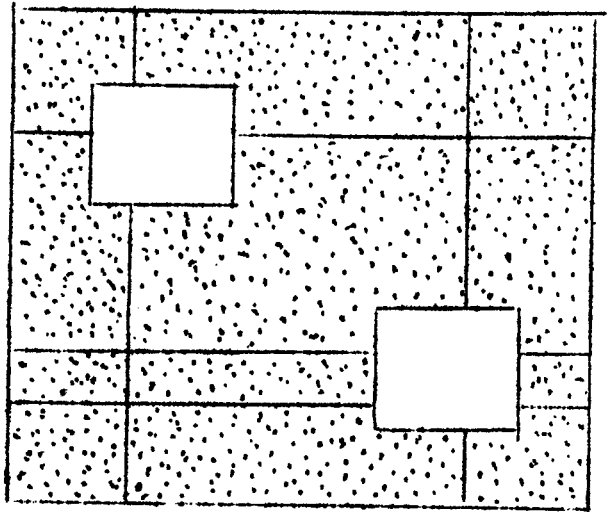
3.



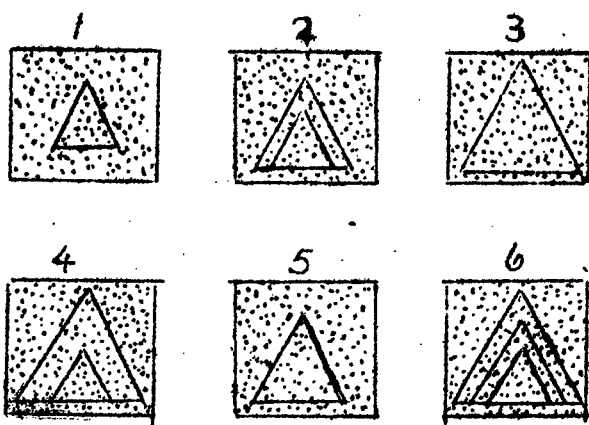
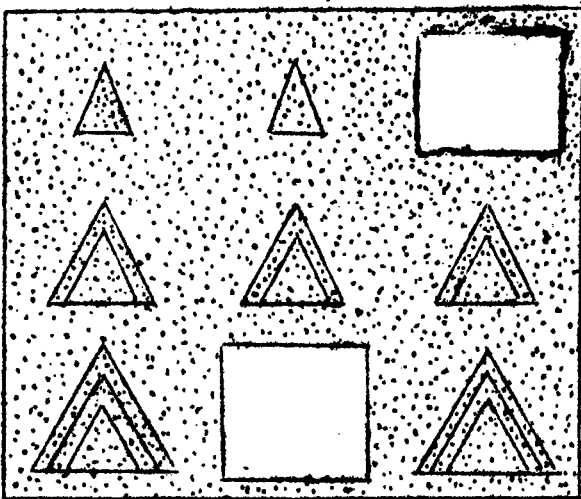
1.



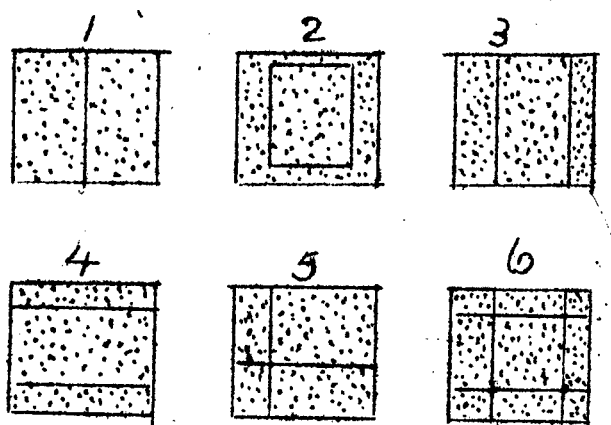
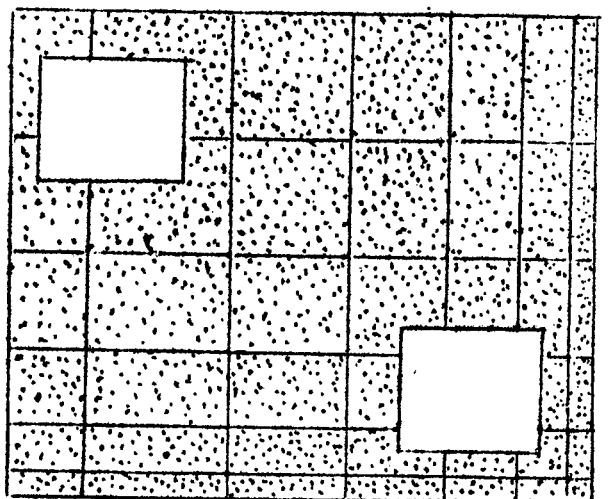
2.

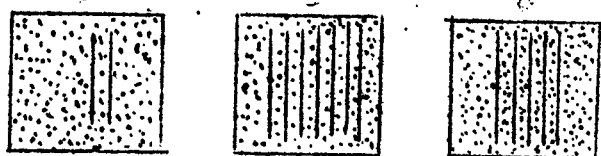
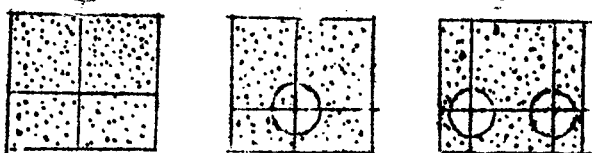
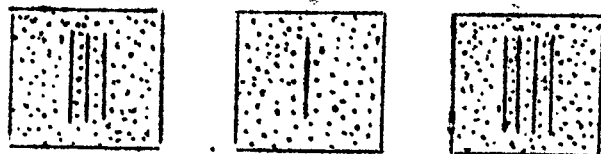
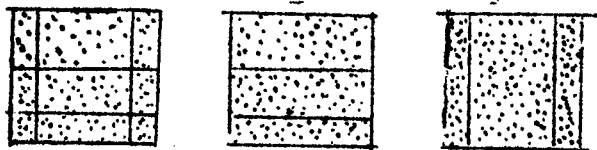
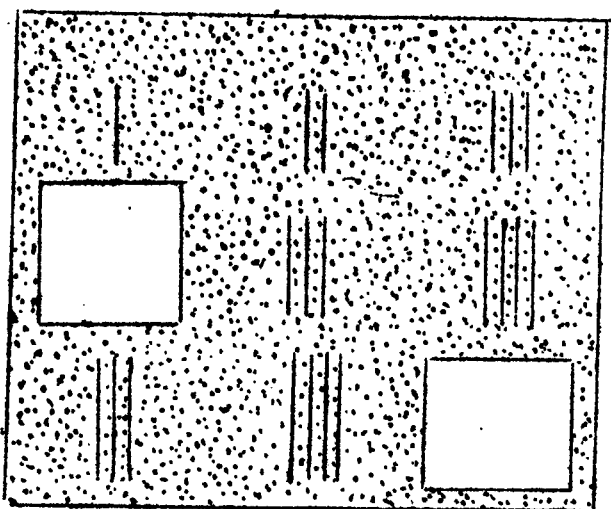
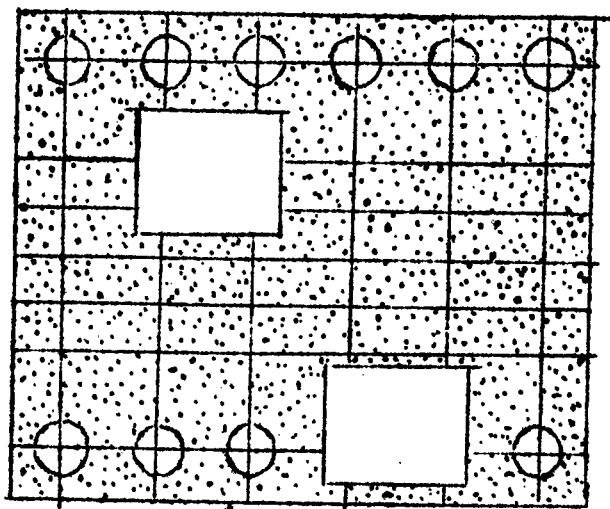
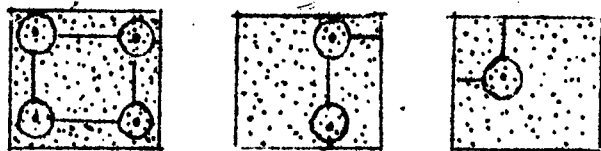
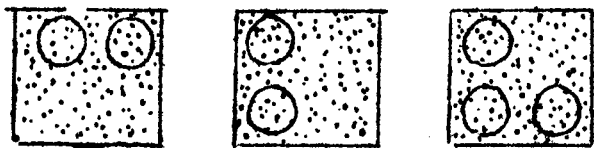
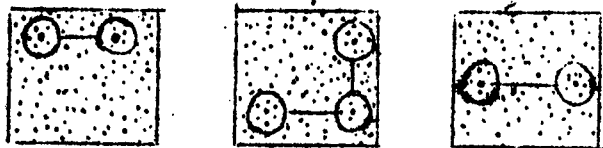
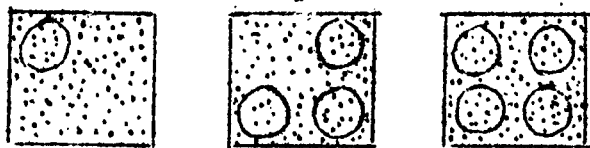
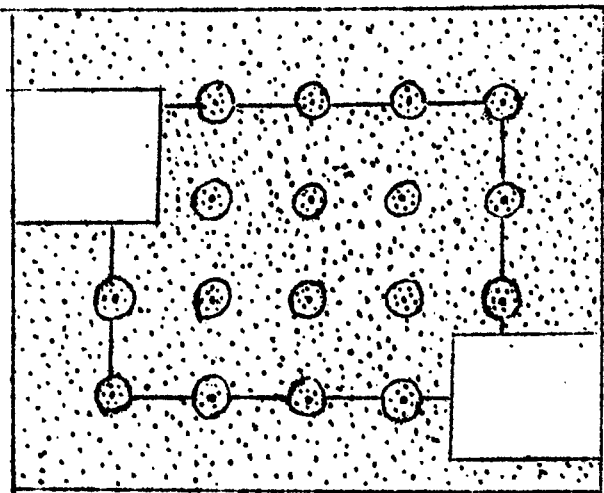
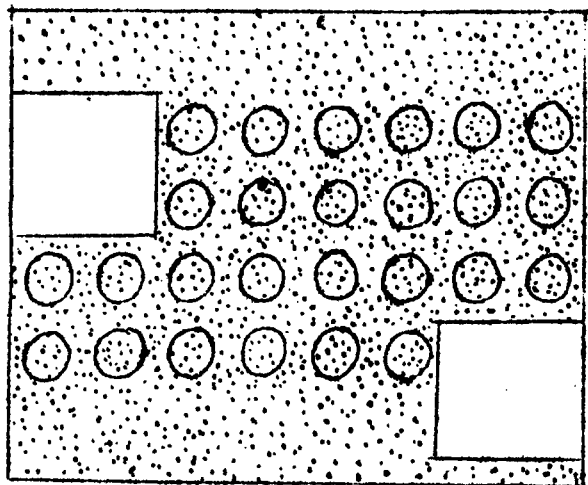


3.

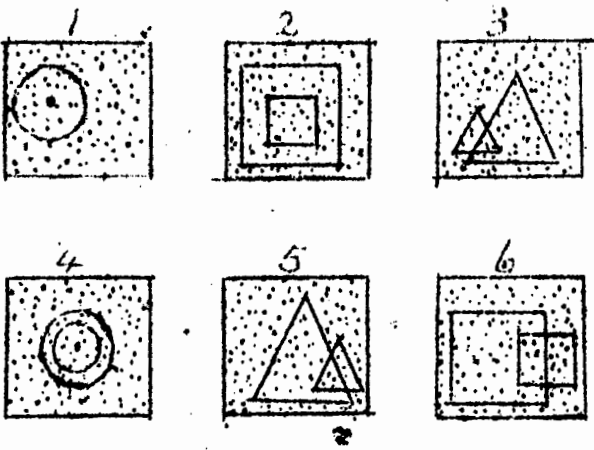
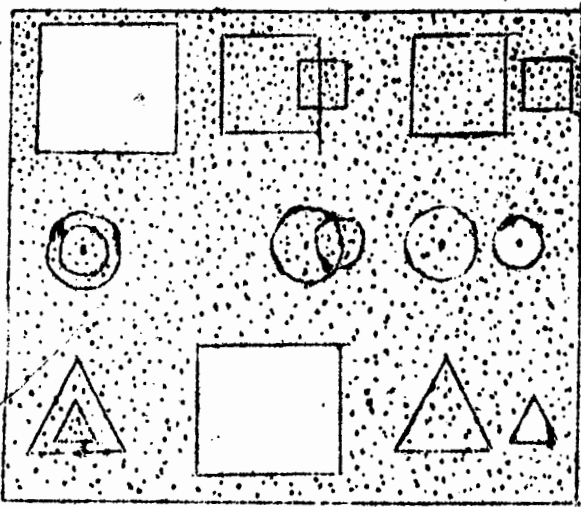


4.

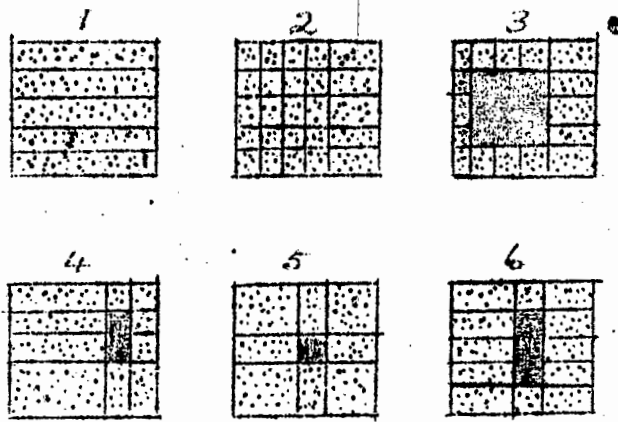
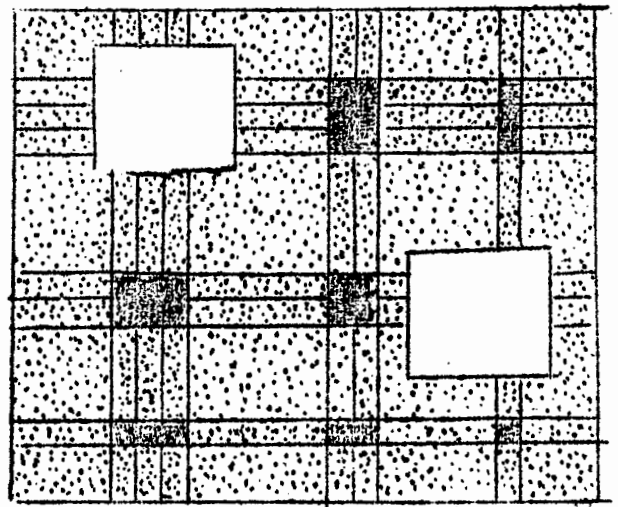




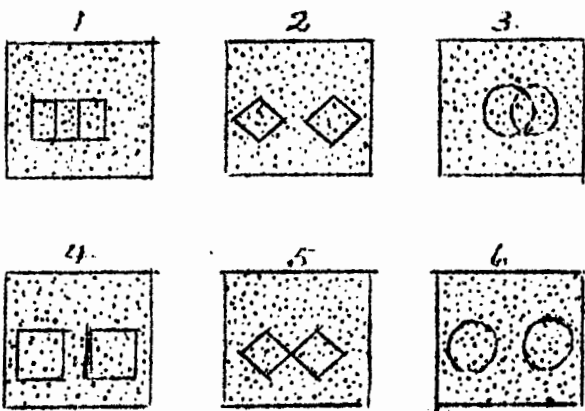
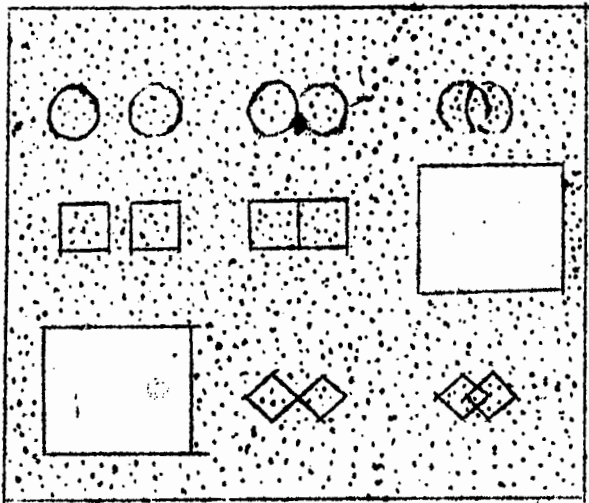
9.



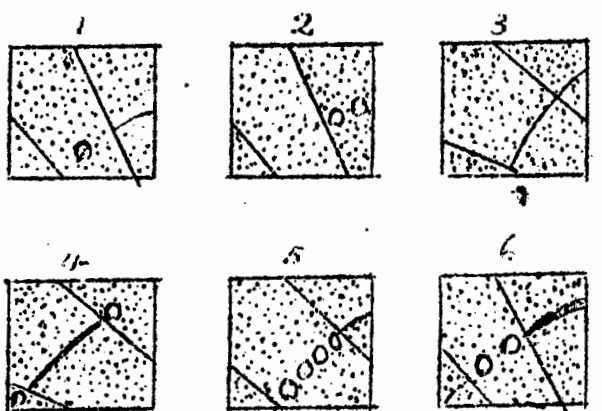
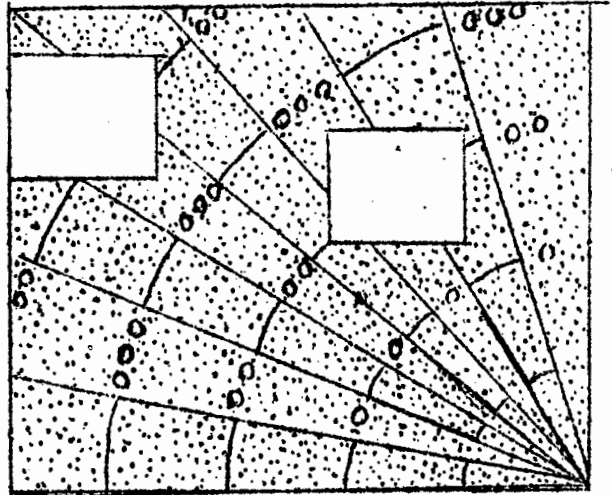
10.

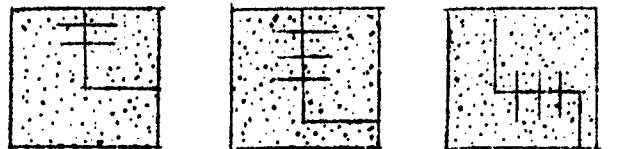
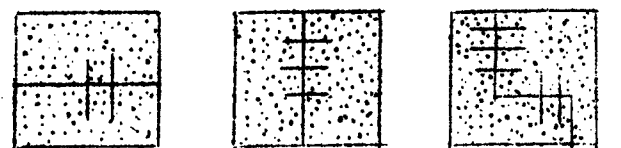
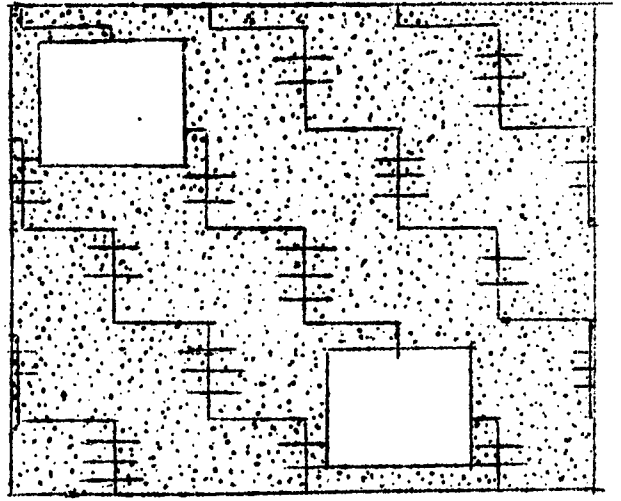
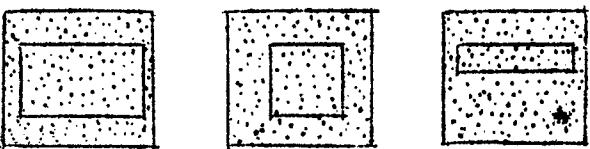
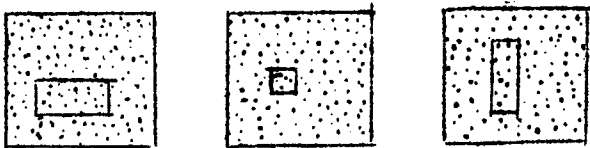
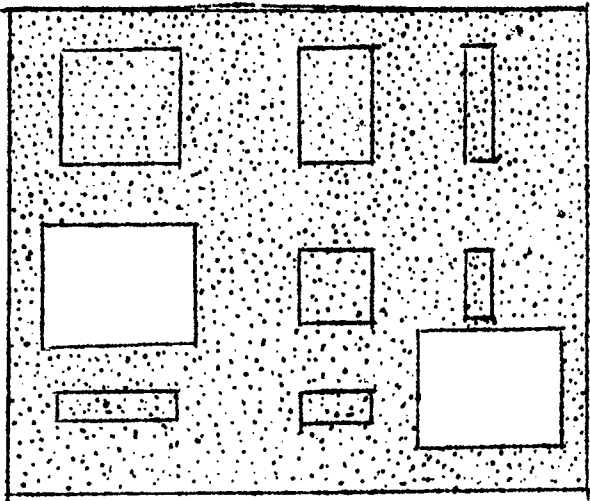
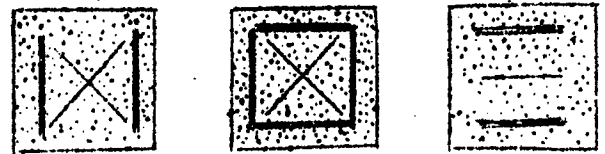
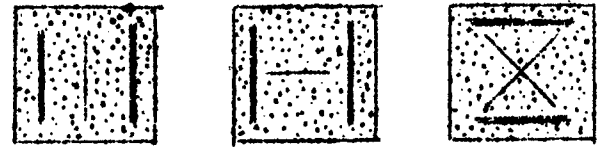
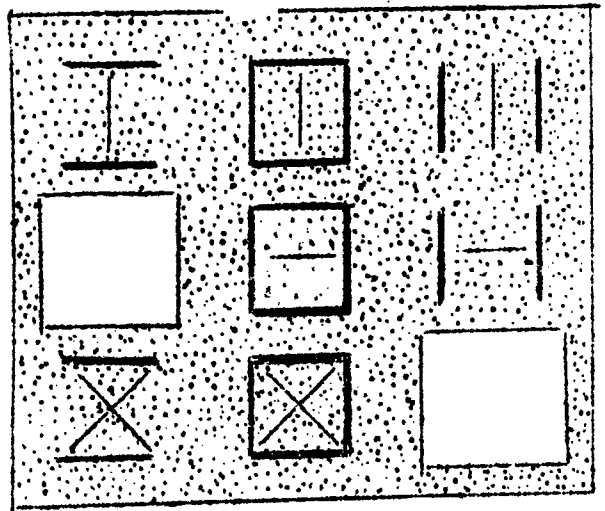
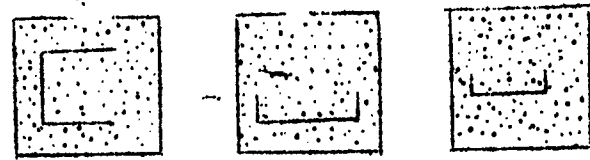
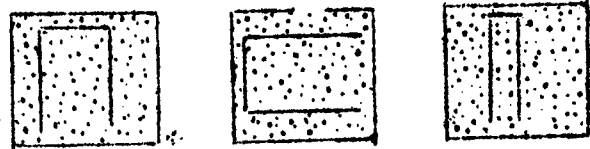
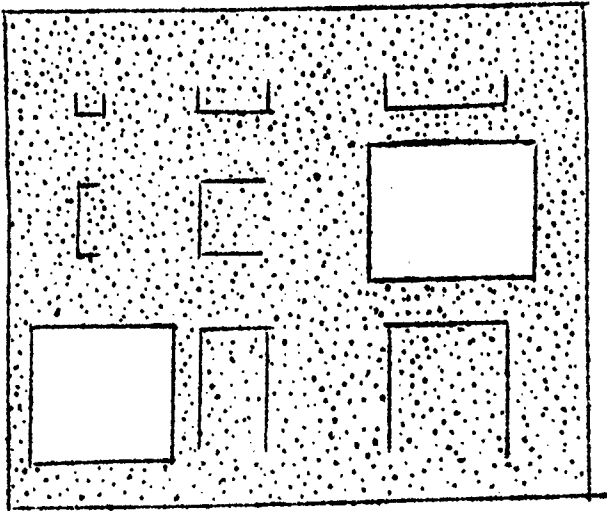


11.

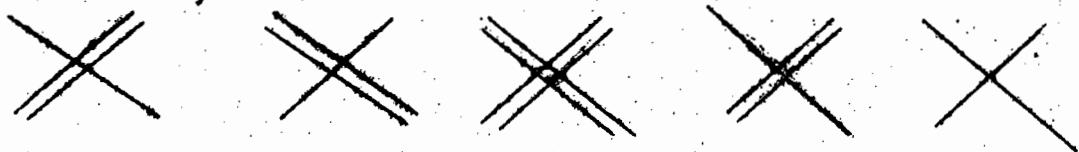
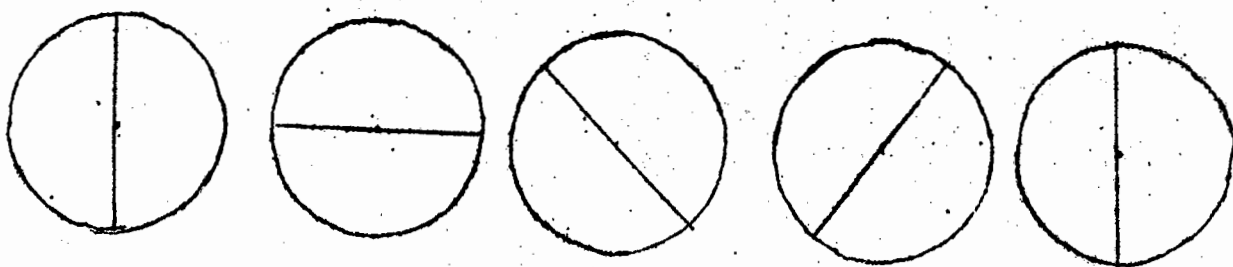


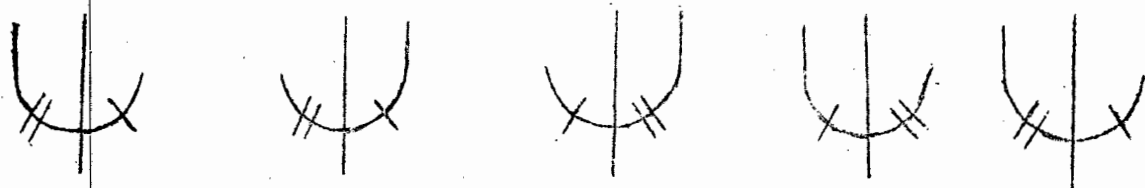
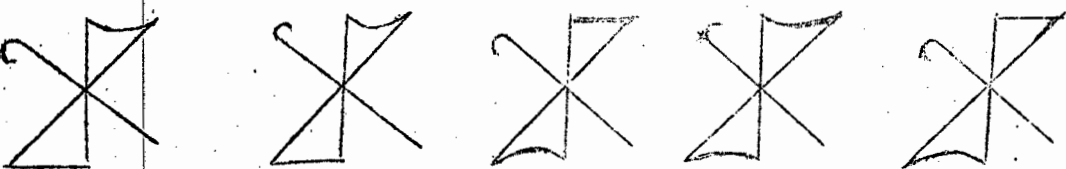
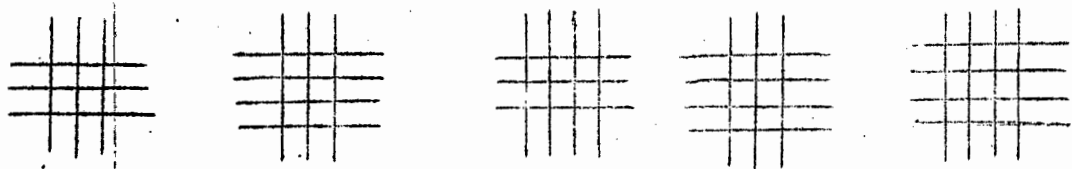
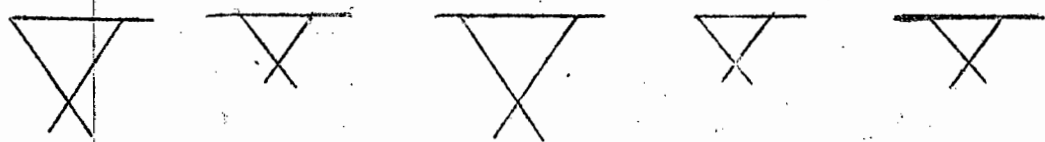
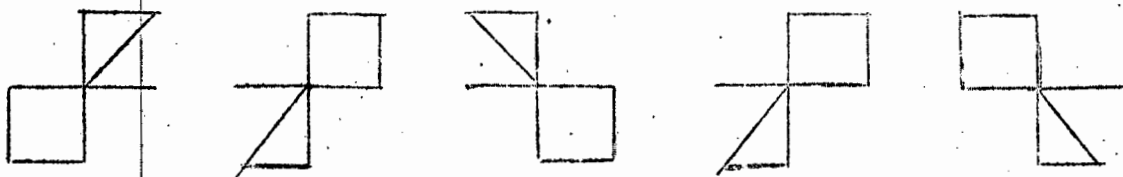
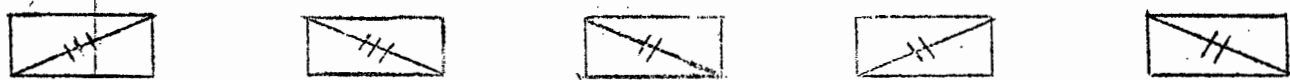
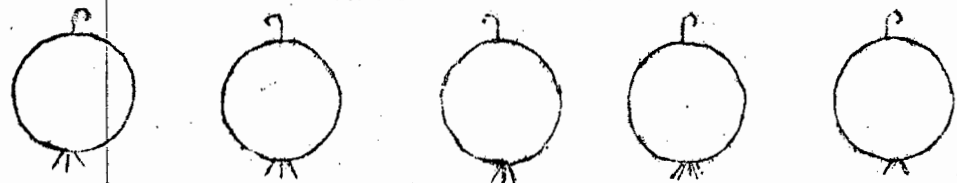
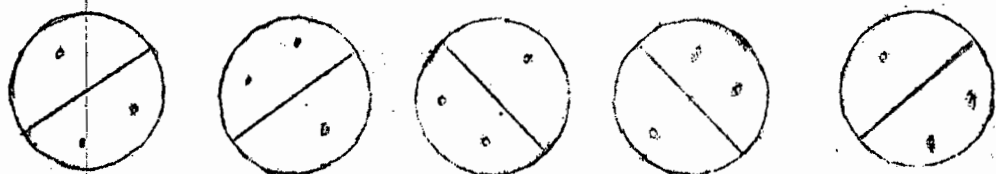
12.

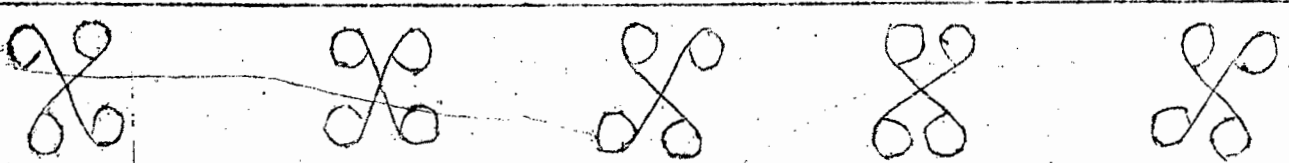
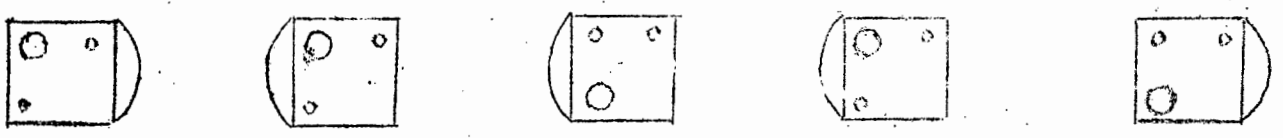
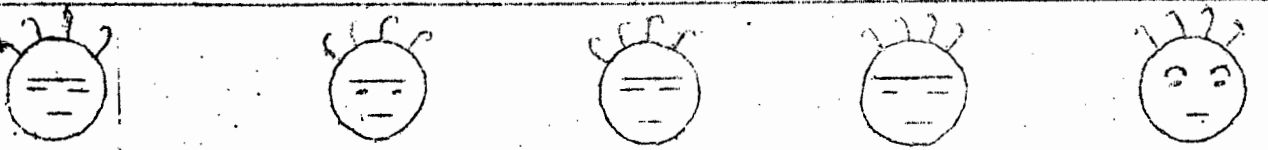
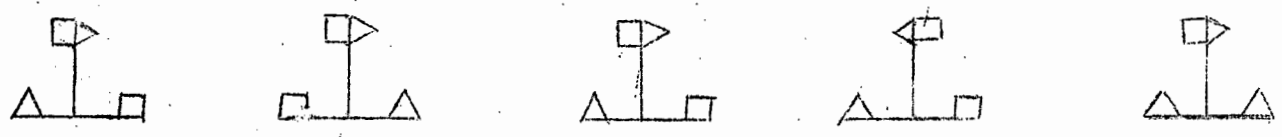
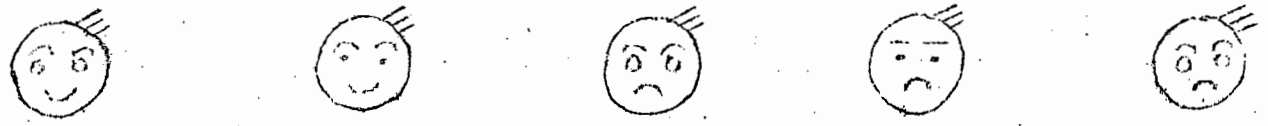
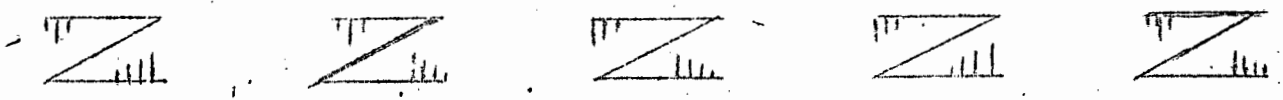
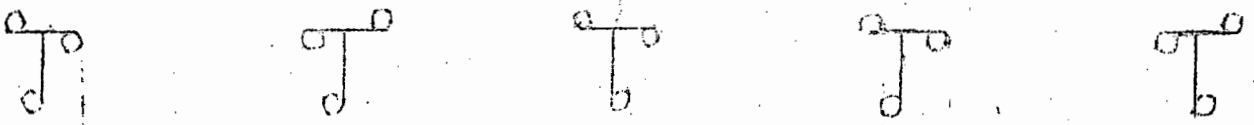
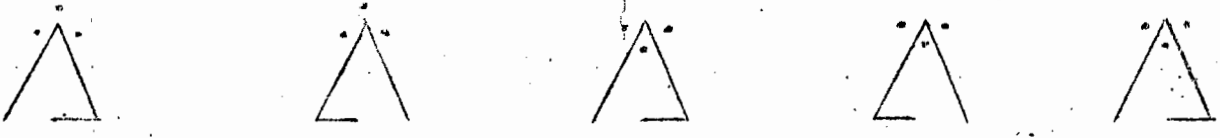


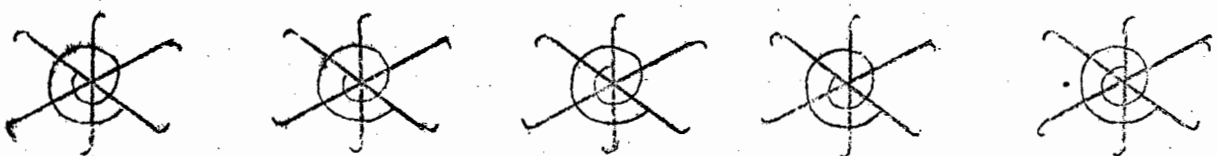
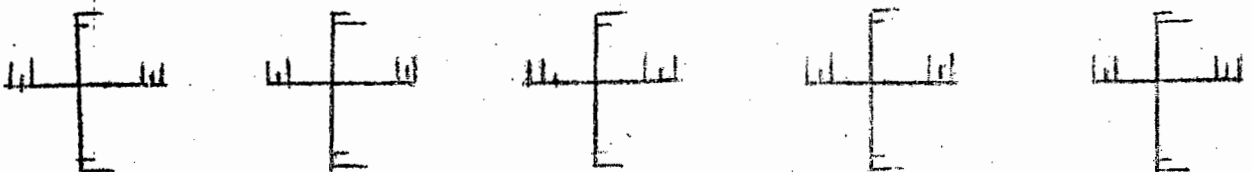
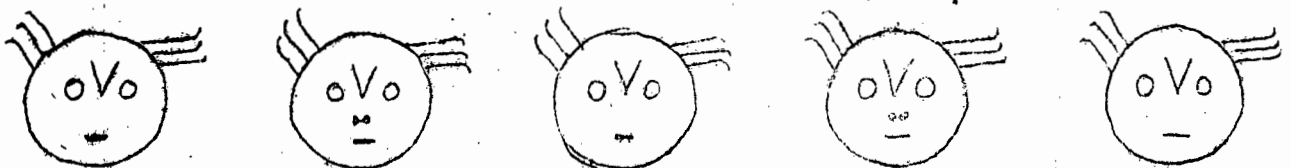
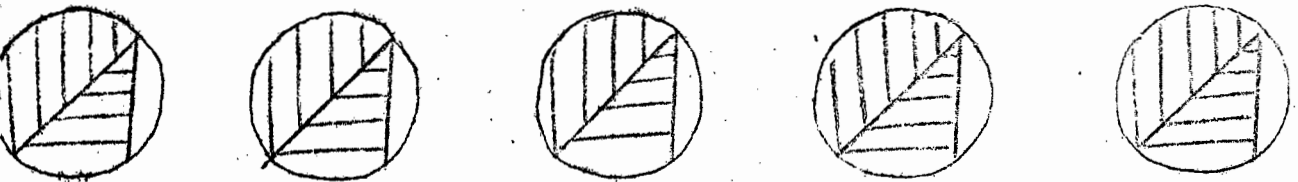
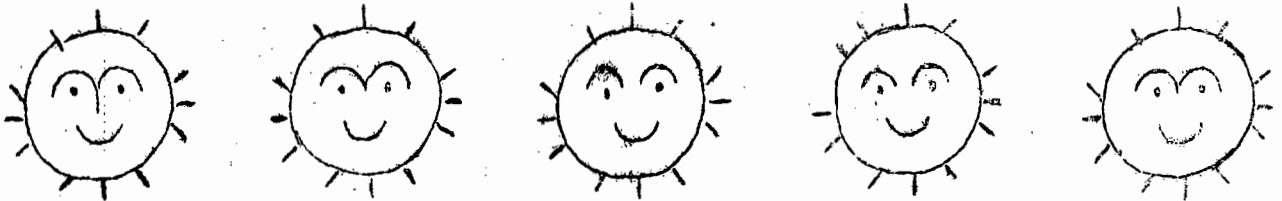
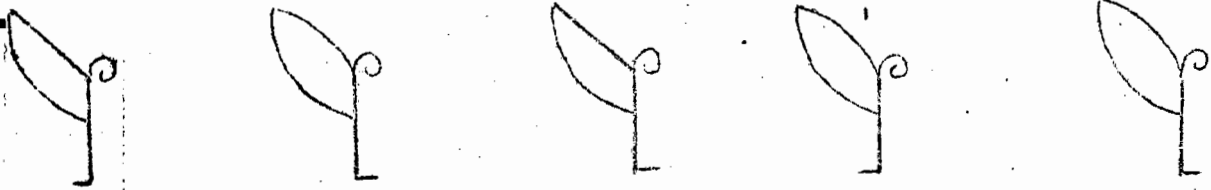
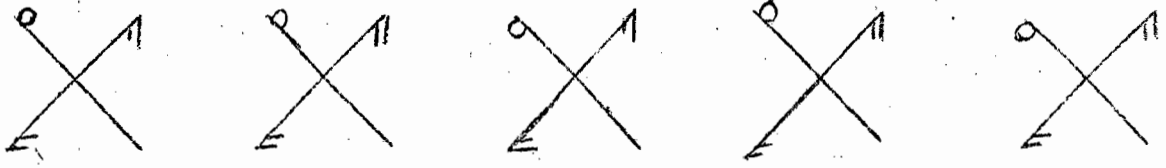


4.

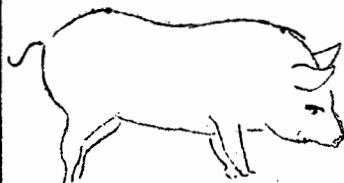
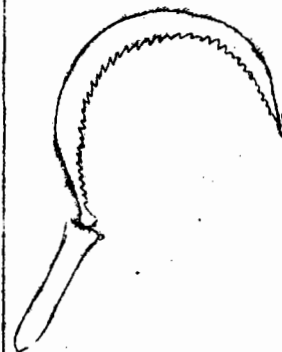
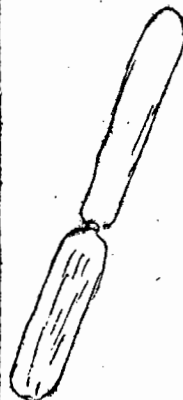
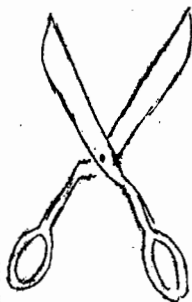
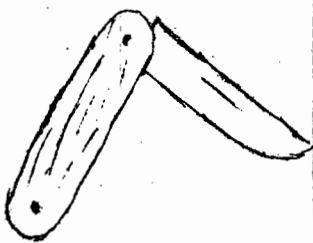
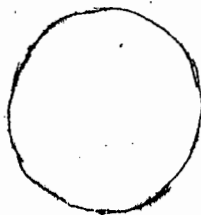
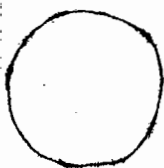


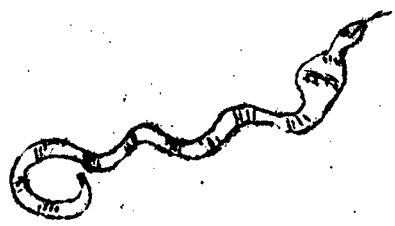
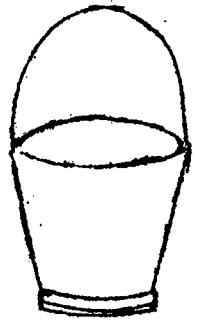
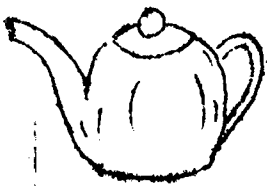
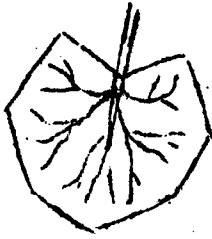
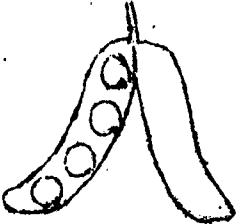
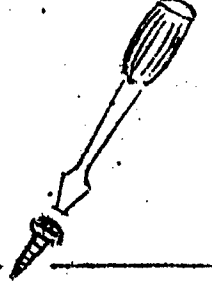
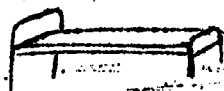
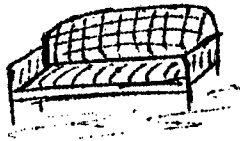
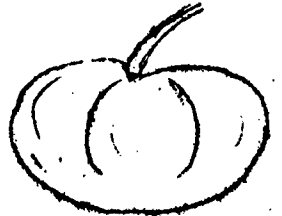
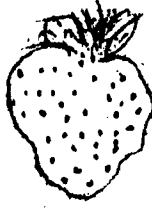
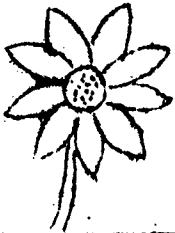
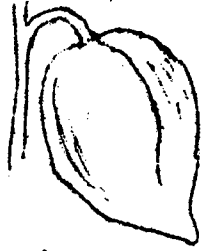
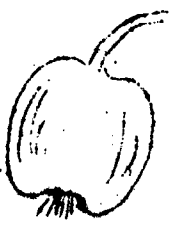
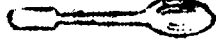
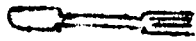
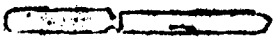
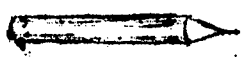
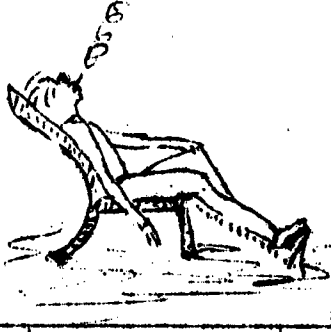
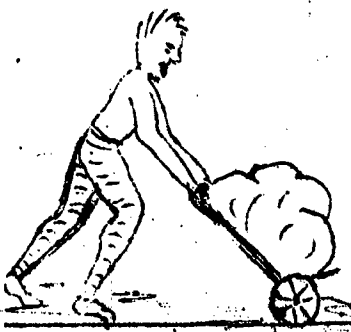


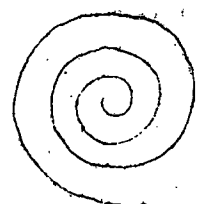
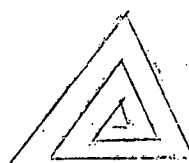
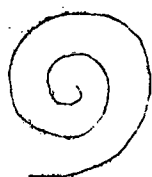
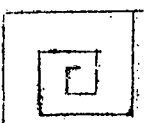
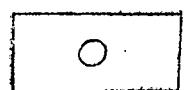
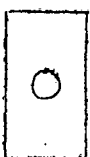
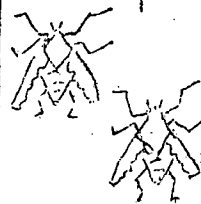
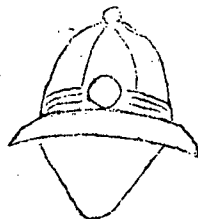
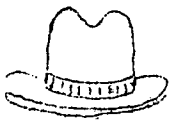
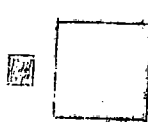
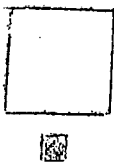
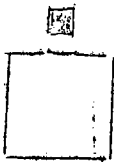
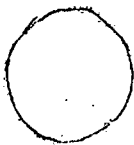
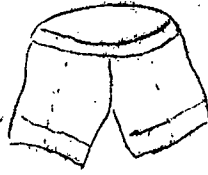
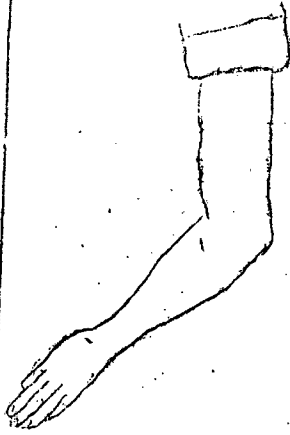
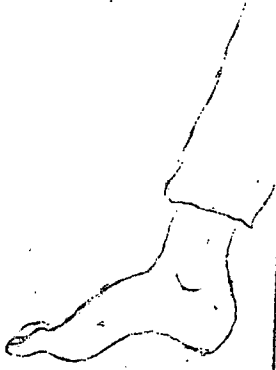


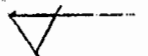
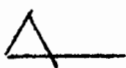
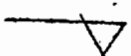
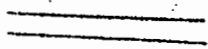
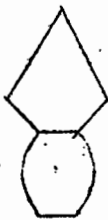
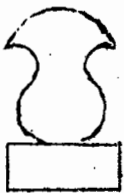
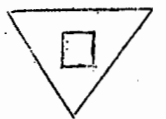
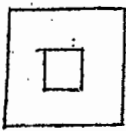
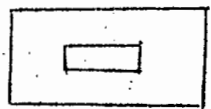
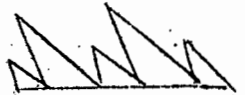
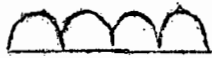
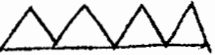
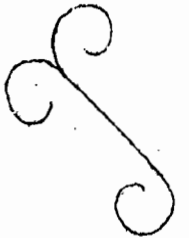
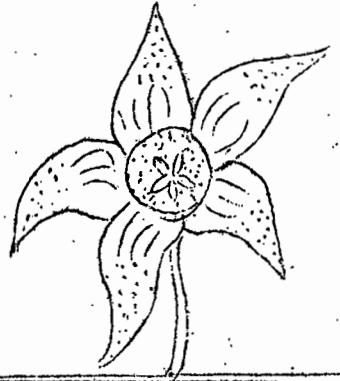
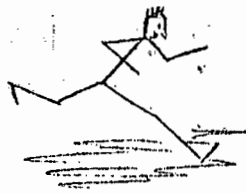
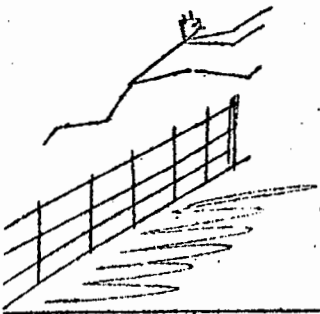


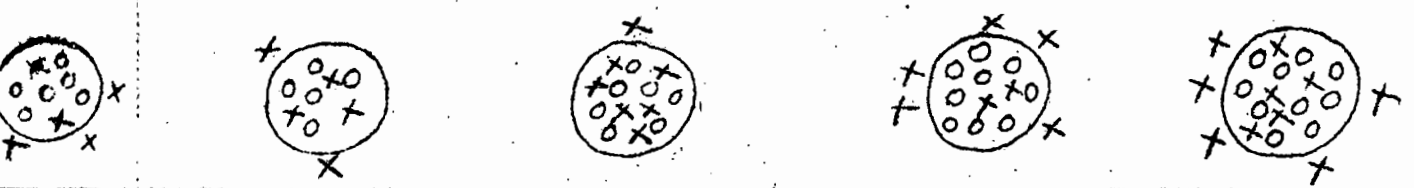
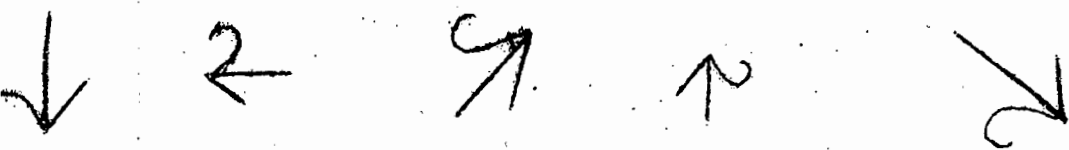
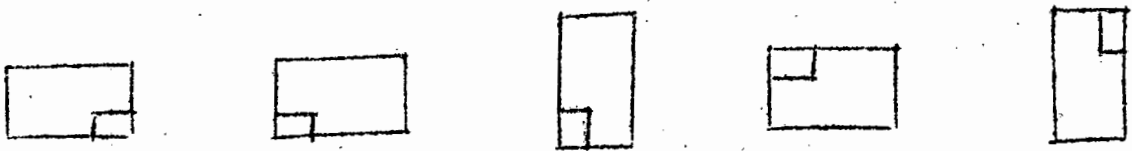
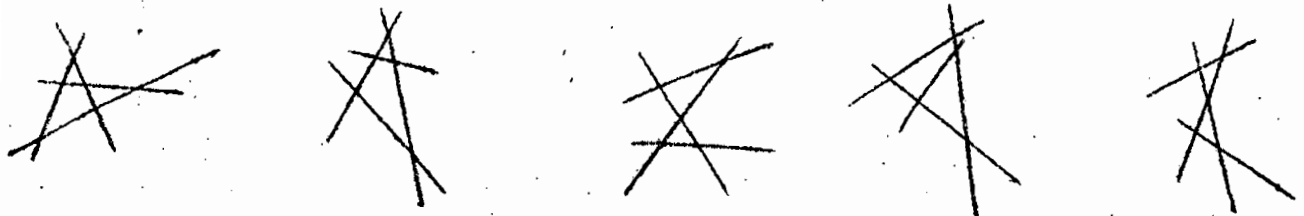
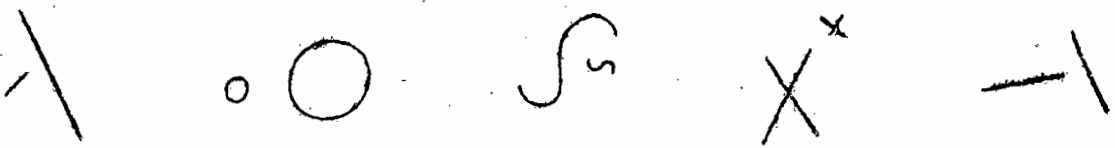
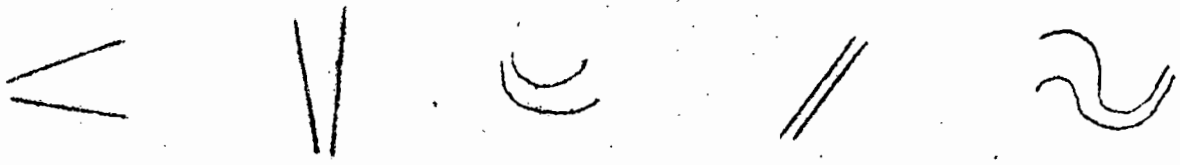
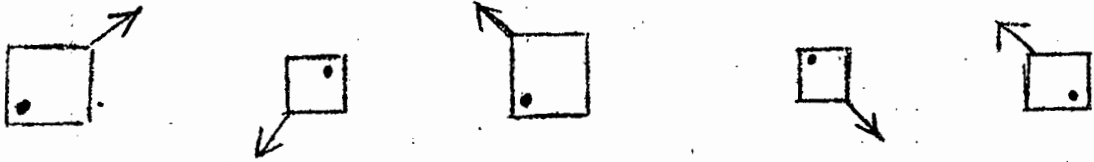
b.











INSTRUCTIONS FOR ADMINISTERING

THE TEST.

Booklets are distributed and the children helped, where necessary, to fill in their names and other details required. This also provides an opportunity to see to it that everyone is furnished with a good pencil, well-sharpened. Rulers are best removed. Erasers should either be provided for everyone individually, or else prohibited altogether. We prefer the latter, as erasing and ruling easily become a most time-consuming nuisance and a highly infectious disease. Provide each child with a loose sheet of clean paper for covering his work.

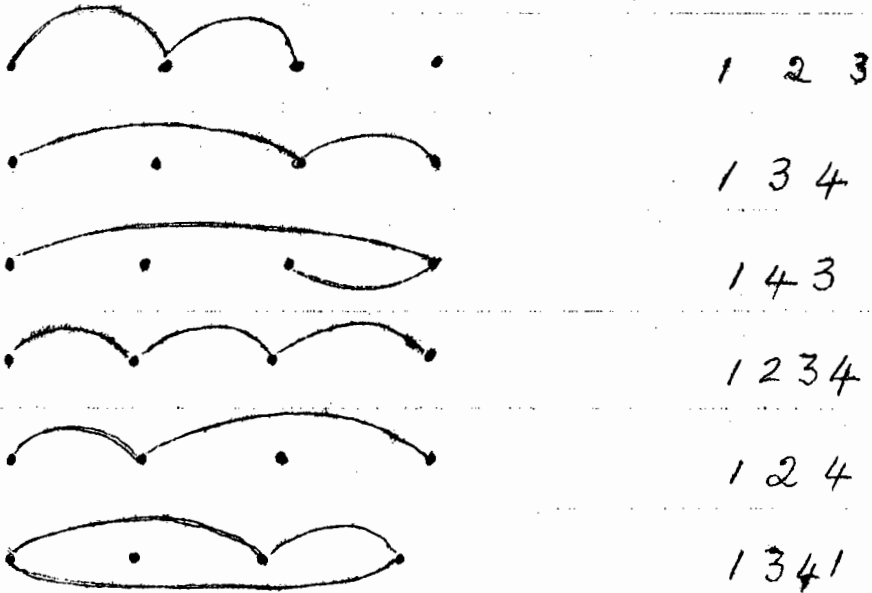
TEST NO. I, FOUR DOTS:

Holding up a booklet, show the children to turn to the page for Test I.

Draw four large dots on the blackboard, about a foot apart. To one side draw several rows of four smaller dots each, to serve as examples. Obtain the full attention of every child. Hold up the blackboard pointer to indicate that it is the instrument about to be used. Standing well to one side trace a curved path from one dot to another with the pointer, pausing a moment on each dot, at the rate of about one per second, in the order indicated in the first example below. On completion, lay down the pointer, and hold up a piece of chalk in its stead to indicate that it is now the chalk's turn to be used. Go to the first row of smaller dots on the side and repeat the movements of the pointer but this time by drawing a clear line from point to point. It does not matter if the reaction at this stage remains blank. Repeat the process for the second and third examples, watching the class carefully the while. By this time the majority would probably have indicated, by facial expression if nothing more, that they have discovered the meaning of what is going on. Proceed to further examples but now call children to the board to do the drawing, being careful specially to call on those who seem to be somewhat uncertain. The examiner must convince himself absolutely that every child understands what is required. If, after going through the six examples provided, there seems to be any uncertainty left, any or all of these examples may be repeated, and if necessary, further easy examples improvised.

...../Examples

EXAMPLES:



When satisfied that everyone understands, erase all examples on the board, leaving only the four large dots for pointing. Now indicate yourself, the pointer and the four dots on the board, and then the children, their pencils and the first line of dots in the booklets, the suggestion being something like: "First I, with this pointer, trace the line on the board; then you, with your pencils, do the same on the first line of dots on your booklets". Repeat if necessary. When going up to the board for the first item, take another look at the class. It will be seen that most of the children have their pencils nicely poised in their hands, ready to do the tracing of their lines simultaneously with the examiner's pointing on the board. Hold up a pencil and lay it down on the table. Repeat if necessary. Indicate that they should all lay down theirs. See that this is done! Point at the class and hold up the flat hand, to indicate that they must wait; then point at yourself to signify that it is your turn first. While doing the tracing on the board, the examiner should stand in such a manner that he can keep his eye on the class. However, once they understand there is rarely any further trouble; although they often forget to lay down their pencils unless they are reminded before every item, they rarely attempt to use them prematurely.

The items of the actual test are given below in numerical form. The dots should be mentally numbered from the left. All left to right motion must be traced above the line of dots, all motion from the right to the left, below. Only touch the board on the dots. The lines between should be traced just off the surface. While the children draw one item, memorize the next by numbers, but be very careful not to give any indication that the dots are numbered.

...../Do

Do not, e.g., allow the lips to shape the words for the numbers soundlessly. They can read your lips! Allow sufficient time for everyone to finish his own tracing. Do not repeat any item. See that each child covers his effort with the loose sheet provided. Before demonstrating the second item, hold up a blank booklet, hold up two fingers, then point to the second line of dots.

If there are more than ten children in the group, assistance will be required. We found one assistant for every ten to fifteen children very useful. Every assistant should move about among the group allocated to him, and see that the children draw on the line corresponding with the number of the item, see that they cover their work, lay down their pencils; but, of course, under no circumstances render any help with the actual task. Any approval asked for should always be given, no matter how wrong the effort may be. Those who hesitate must be encouraged to make an attempt. Never, in any way, indicate that an effort is wrong.

These latter remarks apply to all the tests.

TEST ITEMS:

- | | |
|-----------------|------------------|
| (1) 1 2 3 4 1 | (7) 1 3 2 4 1 |
| (2) 1 4 3 2 | (8) 1 2 3 4 2 1 |
| (3) 1 4 2 3 | (9) 1 3 2 4 2 |
| (4) 1 4 3 2 1 | (10) 1 2 4 3 1 |
| (5) 1 2 3 4 3 1 | (11) 1 3 4 2 1 2 |
| (6) 1 3 4 2 | (12) 1 2 4 3 1 4 |

TEST NO. 2, COMPLETION OF SYMBOL SERIES:

Immediately the first test has been completed, hold a booklet, turn over the page to the practice-page for test number 2 and let the children do the same.

Do the first example of the practice page on the blackboard as follows: first copy the example as it stands; now trace, with the finger, the successive dots and slanting lines slowly; pause at the blank rectangle, and, with a puzzled expression, make a gesture of enquiry; repeat; then, with an air of glad discovery, draw in the

...../missing

missing dot. Indicate to the children that they now have to do the same for the same item on the paper in front of them. Proceed in the same way with the second example. In this case there will probably be volunteers to supply the missing figure. If one fails, let another try. The method of demonstration varies somewhat with the nature of the items. In No. 3, e.g., point to the first rectangle and hold up one finger, to the second, holding up two fingers, and so on. When coming to the blank there is always a pause and the gesture of enquiry. In the case of No. 5, point at the first figure, stand facing the class and stretch out the left arm; then point at the second figure and stretch out the right arm, and so on. Always repeat the performance, once or as many times as may be necessary. Do not, as a result of the many who may have caught the idea and give expression to their eagerness by snapping fingers, jumping up or shouting, lose sight of the one or two who may still be in the dark. In the case of all the items after the first, call on children to come to the board to do them, and as soon as each item has been completed on the board, let everyone do the same item on his booklet. It will be very rarely necessary to do more than four items on the board. Then indicate that they have to proceed on their own. The examiner and assistants now move about amongst the children and should render any help that may be required in the actual solving of items. Do not merely fill in the missing figure causing difficulty, but try to make the child see the principle involved in that special sequence. Items numbers 6 and 7 are the most difficult. After allowing the class to struggle with these for some time, call their attention to the board and solve them as before. In number 6, e.g., lay a pencil on the first arrow, pointing in the same direction, then move the pencil to the next figure giving it a quarter turn, and so on. Now hold up the pencil in front of the class and perform the same turning motion by quarter turns. After, say, three clockwise quarter turns, do one backwards, i.e., anti-clockwise, pause slightly, then shake the head violently to signify that this would be very wrong. Make a fresh start and carry through. Repeat if necessary until fairly general reaction is obtained and a volunteer supplies the correct solution. Better classes may, of course, not find so much difficulty with these two items. In such a case it would be sufficient for the assistants to render individual help to the few weaker ones along the same lines. The last three items are easy and should serve to convince the examiner and assistants that every child has grasped the general idea of the test. However, suggestion may still be offered as far as required.

This practice page is sufficient, if correctly used, to enable every child to obtain a clear notion of the nature of the task. The examiner should be satisfied that this state of affairs has been achieved, then direct the children to turn over the page to the actual test.

...../Holding

Holding up a booklet, point out that they should work down the first column and then go to the second column. They should not work across the page, as the items in the second column are generally more difficult and the weaker child may get stuck too soon and lose heart.

Some sequences contain more than one blank space to be filled in. The children are apt to omit some. The assistants should therefore watch their groups all the time, moving around, and point out such oversights. They may also encourage children and urge them on, signify approval of what has been done if such approval is solicited, but, of course, on no account render any further assistance with the solution.

The next page, containing mostly more difficult items, is intended for possible future incorporation but for the time being serves merely to keep the better pupils, and the quick workers, occupied while the slower ones finish the test page.

We shall discuss the question of time allowance at this point, but it will apply to all the tests, except numbers 1 and 5.

The children should be allowed to work unhampered until only three or four are left who have not finished the actual test page. ^{to} to these and watch their work. If they are still making progress and getting items right, spur them on somewhat but allow them to continue. Very often however, it will be found that such slow ones have reached a stage where they are at a loss. They have gone beyond their depth and by their haphazard attempts give clear evidence that they cannot possibly solve any further items intelligently. As soon as it is clear that the last few have reached the limit of their achievement and are worrying and poring over items they are consistently getting wrong, give them to understand in a kindly way that they have done enough — be careful not to discourage —, obtain the attention of the whole class as always by thumping on the floor, and show them, by means of a booklet held up, to turn over to the practice page of test number 3.

TEST NO. 3 MATRICES:

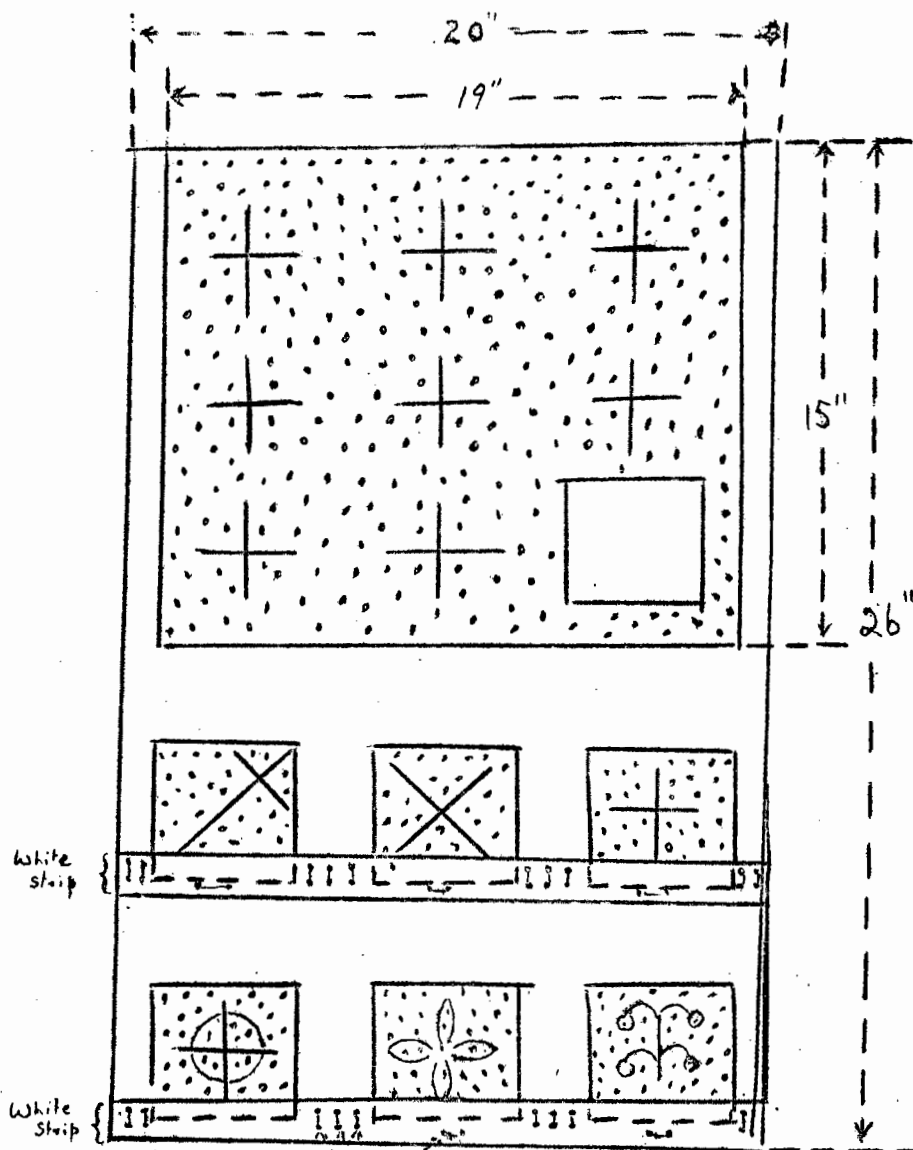
The requirements for the demonstration of this test are as follows:



On a large white cardboard sheet, 20" x 26", another in contrasting colour — we used pink —, corresponding to the first practice pattern, i.e., containing the same figures and a similar rectangular cut-out, is pasted.

...../This

This coloured sheet containing the pattern matrix is 15" x 19", the cut-out being 3½" x 4½". The figures are drawn in very dark blue or black lines, ¼ inch in thickness. The pattern sheet is pasted so that its top edge coincides with the top edge of the white background sheet. The use of different colours for the pattern sheet and the background makes it very clear that there is a rectangular part missing. Six rectangular cards, of the same dimensions as the cut-out and of the same colour as the pattern matrix, are prepared and figures drawn on them as in the booklet. Underneath the pattern sheet, two strips of thin white cardboard are fastened by means of staples or paste, in such a way as to allow the loose rectangular cards to be slipped in and held in the same relative positions as in the booklet, they must, however, not slide in more than about half an inch as the figures might otherwise be obscured.

Two more demonstration charts are prepared, in a similar way, for the other two practice examples in the booklet.



 white
 pink

Staples.

...../Put

Put up the first chart on the blackboard. Hold a booklet, open at the practice page for this test, alongside. Touch the chart and then the figure in the booklet several times to indicate that the two are identical. Lay down the booklet. Trace with the finger slowly the crosses on the pattern matrix, beginning at the top left hand corner. When coming to the cut-out, pause, touch the blank space repeatedly and make a gesture of enquiry. Then take up one of the loose cards provided on the chart - any one except the right one - and fit it into the blank space. Stand back and look at the result appraisingly, then shake the head to signify that it is wrong. Replace this card and repeat with another, giving the children the opportunity to signify satisfaction or disapproval. Use the right card last of all. This will meet with general and probably loud approval. Let one of the children come up to the blackboard with his booklet and pencil. Let him hold his booklet up alongside the chart. Hold up the correct card and touch the corresponding rectangle in the booklet. Repeat several times, to indicate that the two are identical. Now draw with the pencil a thick circle around the rectangle in the booklet, and then once again point out that the figure and the card are similar. Now go through the motion of taking up the card from its original position and placing it in the cut-out blank several times slowly. Pretend to do the same with the rectangle on the booklet. Then draw a pencil line from the circle around the rectangle to the blank, to indicate the movement of the rectangle to the blank space. All the children then have to do this on their own booklets, the one at the board going back to his seat. The assistants move about and help where necessary. Some children will encircle the correct rectangle, but not draw the line to show where it has to go. It is important that this should be done, as the next figures each has more than one cut-out blank.

A similar procedure is followed with the second example and the second chart, except that in this case children are called on to come up to the board and find the correct cards and place them in the right spaces. The assistants once again see to it that everyone does what is required on his booklet and does not omit either encircling lines nor lines of motion.

The children are allowed to attempt the third example on their own, the assistants moving about mainly to convince themselves that everyone understands the nature of the task. They may still help to clear up any individual uncertainties. The third chart is only used in the case of persistent failure to grasp the idea. In such a case the child is brought up to the chart with his booklet where the whole process is very carefully explained once more to him individually by means of the third example, the correspondence of the chart and the figure in his booklet being stressed. If this still does not succeed, the whole test is probably beyond his degree of intelligence and a zero mark on the test will probably be correct for him.

...../Obtain

Obtain the attention of the whole class, let them turn over the page. Point out the blanks and the alternative rectangles provided. Also show them that there are four pages of this kind to work through.

While the children work, the assistants keep a look out for omissions, which are pointed out. Some children will always stop when they come to the end of a page. In such a case the assistant should turn the page for him and tell him to proceed by pointing out the further items.

All four pages are part of the test.

After this test an interval of at least fifteen minutes should be allowed.

TEST NO. 4. IDENTICAL PAIRS:

This test is given immediately after the interval

The first example of the practice page of this test is drawn on the blackboard, the children having turned over to this page.

Trace each of the figures slowly with the finger. Repeat for the two identical figures. Tap them simultaneously with the two hands. Then draw a large, bold circle around each of the two. Let the children do the same on their booklets.

Draw the figures for the second example on the board. Tap the first two figures and look enquiringly at the class. Some of them are sure to express their disapproval. Repeat for the first and third, the first and fourth, the first and last, the second and third, the third and fourth figures, shaking the head with each combination. By this time some one will probably be clamouring for the opportunity to come and help the examiner out of his difficulty. Give him the chance. Let everyone do this example on his booklet.

Put the third example on the board. In this case ask for volunteers from the start. Call on someone who does not volunteer. If he chooses the wrong two figures, point out the differences and let him try again until he succeeds.

Let the children attempt the last two examples on their own. Assistants move around and help, mainly by pointing out differences where wrong pairs are selected. In cases where more than two figures are encircled, they should also give the child clearly to understand that only two figures should be chosen.

These last two examples should serve to show whether everyone understands the requirements of the test and to clear up any remaining uncertainties. When this has been achieved to the examiner's satisfaction, let them proceed to the actual test, but first hold up three fingers and shake the head negatively, repeat with four fingers, then with five. Then hold up two fingers, nodding affirmatively. Repeat the whole process.

While the children work at the test, the assistants should be on the lookout for those who still mark more than two figures in one row. Point this out to the child and explain once more that two, and two only, should be marked.

This may be done a second time for the same child if necessary. If the child continues making the same mistake after being shown twice, it would be of no avail to go on explaining and he should be left to his own devices. Note that it is only the error of marking more than two figures that may be pointed out. No other errors may be indicated in any way. Attention should also be called, as in all tests, to omissions of items.

The first two pages constitute the test, consisting of 25 items. The last page is again intended to keep the fast workers employed while the slower ones finish the first two pages.

As regards time allowance the same procedure should be adopted as explained in connection with test number 2.

TEST NO. 5. FIGURE RECALL:

R e q u i r e m e n t s: Ten cardboard sheets, 8 inches by 10 inches, white or some light colour. On these are drawn in black or dark blue, lines $\frac{1}{4}$ inch thick, the geometrical figures, reproduced on the next page, dimensions to be three times those of the reproductions here given.

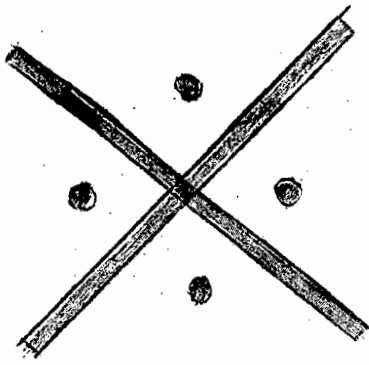
The children are shown to turn over to the page in their booklets left blank for test No. 4. This page has been marked off into ten rectangular spaces. They must then lay down their pencils and attend for demonstration and explanation.

The first two items are used for this purpose. Hold up the first card for about five seconds. Then lay it down quickly, go up to the board and draw the same figure, omitting one of the four dots. Look enquiringly at the class as if requesting their approval. Some pupils are sure to notice the omission immediately. Let one of them come up and draw in the missing dot. Now take up the card again and hold it immediately above or below the drawing, trace the figure with the finger first on the card, then on the board, smile with satisfaction and nod emphatically to indicate that the drawing on the board is correct.

The second card, containing two figures, is used as a further example. Obtain attention, hold up the card for seven seconds, then call on one of the pupils to come to the board and draw the figures. Probably the first child will not succeed, but even if he does not make a single mistake, pretend not to be quite satisfied and call on someone else; and then on two or three more. Leave all attempts on the board. Now take up the card again and hold it next to each attempt and let the children express in their own way which attempts they see to be successful. We have not found it necessary to give any further practice or explanation of the general nature of the test.

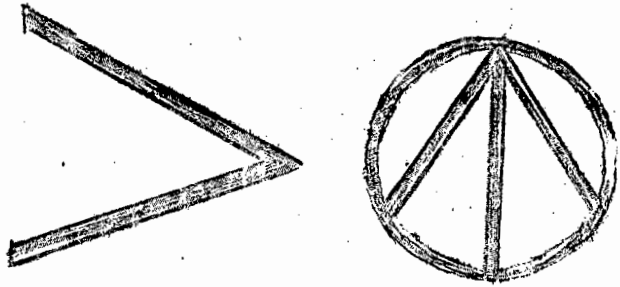
The examiner now points at himself, then pretends to take up a card from the pile on the table and to hold it up, then to lay it down with a quick motion. Then point at

Demonstration

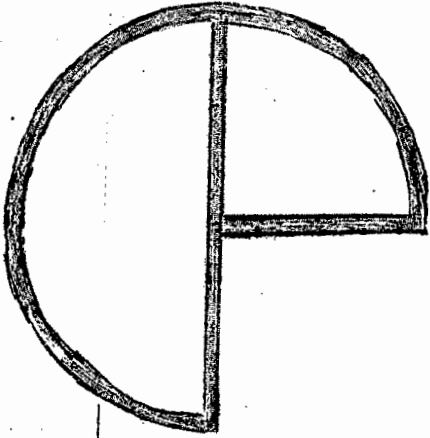


No. 1. (5 secs)

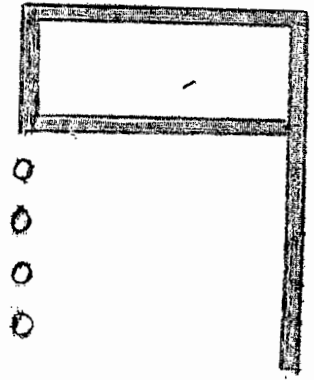
Demonstration.



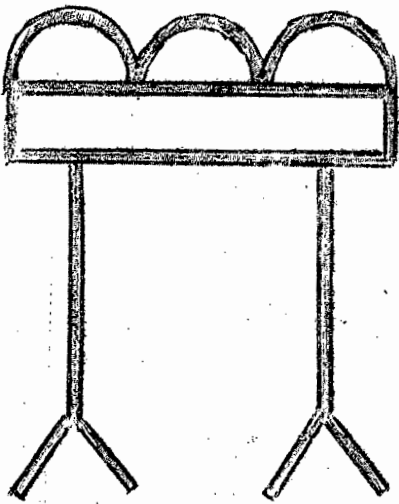
No. 2. (5 secs)



No. 3 (5 secs)

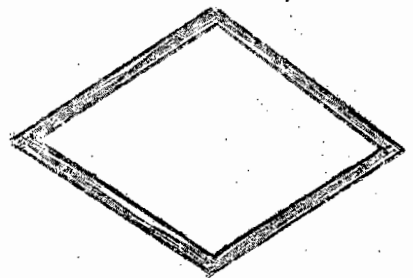
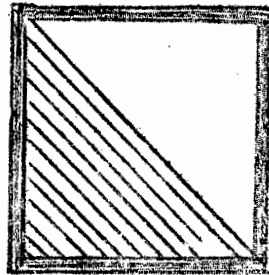
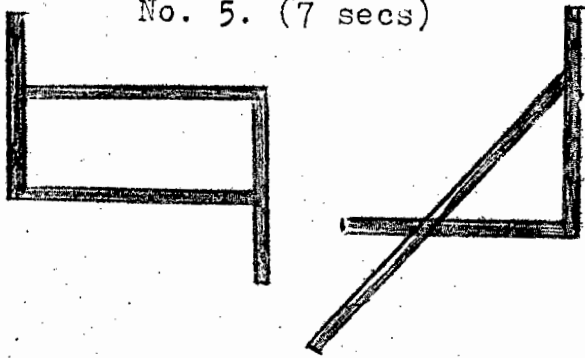


No. 4. (7 secs)



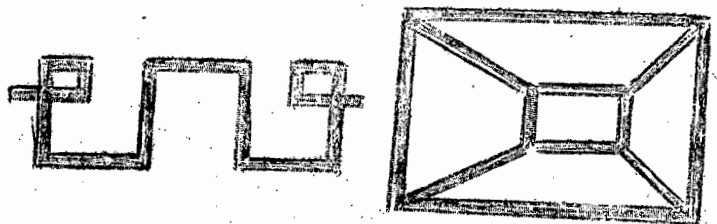
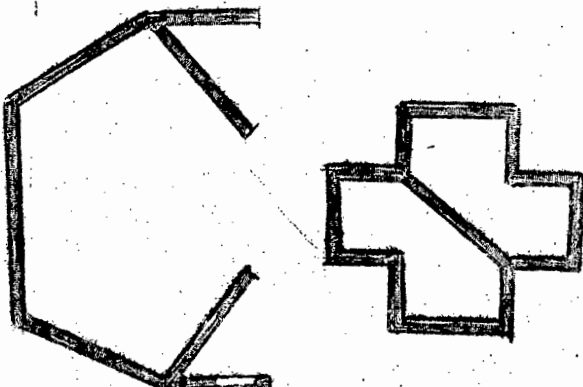
No. 5. (7 secs)

No. 6. (7 secs)



No. 7. (7 secs)

No. 8. (10 secs)



the class, hold up a booklet and with a pencil make a drawing movement in the top left hand rectangle. See that the examples on the board are erased.

Obtain the attention of everyone. Let them lay down their pencils. Show the two demonstration-cards briefly again, shake the head and signify with the hand that they have been finished. Lay them conspicuously apart, take up the first of the test cards, hold it backwards so as not to reveal the design, tap it and nod in acquiescence and point again at the booklets and the pencils and the class.

Call everyone to the alert once more, twist the first card, still being held in the hand, around, hold it up, making sure that everyone can see the design perfectly, and after exactly five seconds, lay it down quickly, face downwards and again point at the class and the booklets. The assistants should now see to it that the children all make a start and draw in the correct space. They should also cover their work immediately.

Before exhibiting the second card, hold up a booklet once more and point at the top right-hand rectangle where the second drawing should be made.

Proceed in a similar way with all the cards. Cards numbers 1, 2 and 3 should be shown for 5 seconds, numbers 4, 5, 6 and 7 for 7 seconds each, and the last one for 10 seconds.

Where more than one figure appears on a card both should be drawn in the same space in the booklet. The assistants should see to this. The two additional spaces are provided in case some children may have already erred in this respect before the assistants could prevent it. In such cases allow the drawings to stand, but point out to the child that in the next case he should draw the two figures in one space.

TEST No. 6. CLASSIFICATION:1). Let the class turn to the practice page for this test.

While the children were finishing the previous test, the first three examples of the new test should have been roughly sketched on the blackboard.

Obtain attention. Point at each of the fruits represented in the first example in turn, and in each case bring the hand from the drawing to the mouth, making an eating or chewing motion with the mouth and jaws.

- 1). We owe items numbers 14, 15 and 17 of this test to a similar test by Oliver in his "General Intelligence Test for Africans", by kind permission of the Director of Education, Kenia (Education Department, Nairobi.)

Then point at the tree, again bring the hand to the mouth but pause, with an enquiring expression. When the children signify disapproval, cross out the tree. Let everyone do the same on his booklet. The assistants should see that this is done.

Turn to the second example. Point at the representations of grown-up people, after each holding the hand high above the floor to indicate size. Then point at the representation of the child, lowering the hand, to indicate smallness. Repeat. Then cross out the child. Let the class do the same.

In the case of the third example, point at each circle, at the same time making a circling motion with the other hand. When coming to the square do the same, then pause enquiringly, and repeat this for the square. The children will say "no" in no uncertain fashion. Cross out the square and let them do the same on their booklets. The remaining three examples have to be tackled by the children on their own, with the help of the assistants where necessary. The help offered should be suggestive rather than explicit. The aim should be to help the child to find the solution himself. E.g., in case of the fourth example, the assistant points at the birds, making a flapping motion with his arms to indicate flight, then at the rabbit enquiringly. These last three examples should again serve to show whether the children have grasped the idea of the test, and to clear up remaining uncertainties, such as marking more than one figure in each row.

Before proceeding to the test itself, hold up two fingers, then three, then four, shaking the head negatively. Then hold up one finger, nodding affirmatively and emphasize by repetition.

As regards this error of marking more than is required the procedure during the test is exactly similar to that in case of the test number 4, except that the assistants here point out, if necessary in the case of two items for the same child, that only one figure should be marked in a row. For the rest the functions of the assistants are the same as before. See that each page is turned as soon as completed and that work proceeds on the next.

The test consists of three pages, the fourth being--again intended to keep the fast workers employed while the others are finishing. Time allotment as before.

INSTRUCTIONS FOR ADMINISTERING THE TESTS TO
HEARING SUBJECTS:

The method of presentation should in essence remain unaltered.

A purely pantomimic performance by a hearing examiner testing hearing children may easily cause the situation to become strained, unnatural and awkward, everyone feeling slightly ridiculous. For this reason the employment of speech is desirable and even necessary. The exact words to use have not been rigorously prescribed so that the examiner may express himself in his own way, but the general trend should be as follows and should, as regards content, be confined to this: After telling the children to fill in their names and the other particulars required, proceed somewhat as follows: "We have some interesting problems here. I would like to see how well you can solve them. Do as many as you can and try to work quickly. Some are quite easy, others are more difficult, but just do your best."

"I am not going to tell you what to do. Instead, I want you to watch carefully what I do on the board by way of examples, and see if you can find out for yourselves what you have to do. Now look! (pointing the first sequence of the four dots-test), and now (tracing the same sequence with chalk). And here is another example (proceed as before). Now I shall give one of you a chance to show us what he can do." And so on.

It will be found quite easy to do as much talking as may be found necessary to put everyone at ease without doing any actual verbal explanation of the requirements of the tests.

When using the tests on a language-group whose language is not understood or spoken by the examiner the best method is that prescribed for the American Army Performance Scale Examination for a similar case which, slightly adapted, reads as follows:

"The Examiner should take care that his directions do not appear too artificial. For this reason he should not always remain absolutely silent. He should try to use whatever words are intelligible to his subjects: 'No', 'Yes', 'Hurry' etc., can be used in most cases, and even when the subjects do not understand, it is often better for the Examiner to speak as well as gesture. The aim here is only to make the instructions intelligible apart from the language used."²⁾

2) Yoakum and Yerkes: "Mental Tests in the American Army", London, 1920.

INSTRUCTIONS FOR SCORING.

Scoring is very simple and straightforward in the case of most of the tests. The following remarks will help to ensure objective marking.

TEST NO. 1:



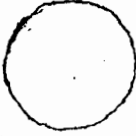
One mark for every correct item.

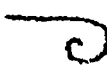






Possible total: 12.





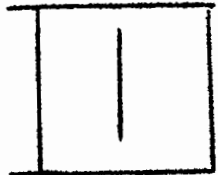



TEST NO. 2:

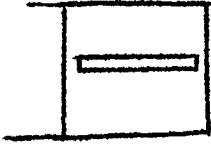

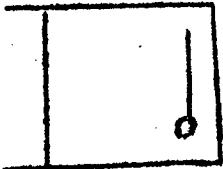
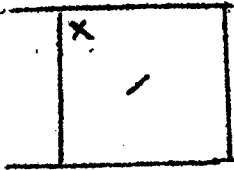
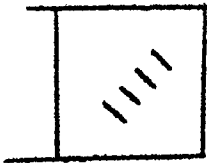
One mark for every correct item.









Possible total 33.

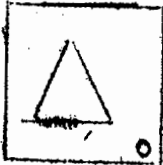
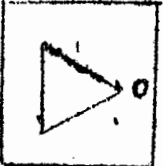


ITEM NO:	CORRECT RESPONSE:	REMARKS:
1.		The line should be noticeably shorter than the one immediately preceding it.
2.		
3.		The circle should be noticeably larger than the preceding one.

ITEM NO:	CORRECT RESPONSE.	REMARKS.
4.		
5.		
6. 1st blank space.		
2nd blank space		
7.		
8.		
9.		<p>The three dots may be placed in any position.</p>

ITEM NO.	CORRECT RESPONSE	REMARKS:
10.		Circle to be noticeably larger than the next one.
11.		The two dots may be placed in any position.
12. 1st blank		
2nd blank		
13.		
14.		Anyone of these alternatives.
15. 1st blank		
2nd blank		

ITEM NO:	CORRECT RESPONSE:	REMARKS:
16.		<p>The rectangle should be both longer, and less in height than the preceding one, but credit is also given if it is at least equal in length but less in height or at most equal in height but greater in length than the preceding one.</p>
17.		
18.		<p>The placing should be noticeably to the right of the centre of the space.</p>
19.		<p>The line should be noticeably shorter than the preceding one.</p>
20.		<p>The four lines may have any length and may be placed in any relative position. The only essential is that there should be four. Four dots and even the figure 4 are accepted.</p>

ITEM NO:	CORRECT RESPONSE:	REMARKS:
21. 1st space		
2nd space		
22.		The dots may be placed in any relative position, or the idea 5 expressed in any other way.
	or 	
	or <i> etc.</i>	
23. 1st space		
2nd space		Some children attempt to show that there are two lines by drawing both lines not quite coinciding, as in the second example. This is accepted, if they are close together.
	or 	
24.		The line should be noticeably longer than either of the two preceding lines though not necessarily equal in length to those in spaces 1 and 4. The slant should be definitely from the top left to the bottom right though not necessarily pointing straight at the corners. The dot should be noticeably smaller than either of the two preceding circles, and filled in, not an empty circle.



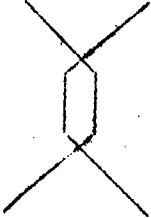

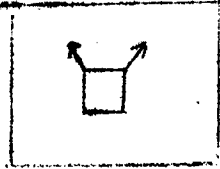



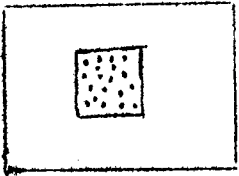
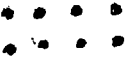
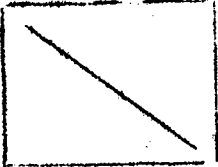




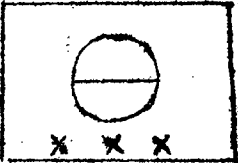
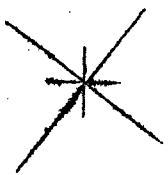
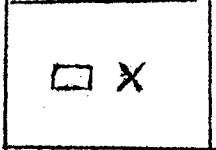
ITEM NO:	CORRECT RESPONSE:	REMARKS:
25. 1st space		The correct placing of both the triangle and the small circle is essential
2nd space		
26. 1st space		
2nd space		

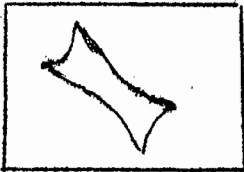


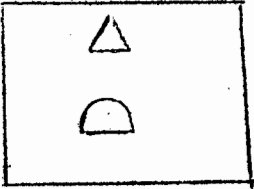
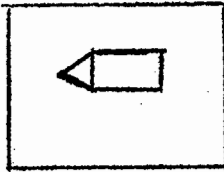
The above items constitute the test. The correct responses on the items on the next page of this test are also given below. Marks obtained on these are to be disregarded entirely and should under no circumstances be added to those obtained on the test proper.

...../Correct Solutions

Correct Solutions of Additional Items, not Part of the Test.

No Marks to be Allotted!

ITEM NO.	SOLUTION.	ITEM NO.	SOLUTION.	ITEM NO.	SOLUTION.
1.		6.		12.	
2. 1st space		7.		13.	
2nd space		8.		14.	
3.	 <i>(any relative position.)</i>	9.		15.	
4.		10.		16.	
5.		11.		17.	

ITEM NO.	SOLUTION.	ITEM NO.	SOLUTION.	ITEM NO.	SOLUTION.
18.		20.		22.	
19.		21.			

TEST NO. 3.

The required rectangle and the cut-out space in the matrix for which it is intended should both be indicated correctly to obtain credit of one mark.

Possible total: 32.

When the correct rectangle is marked but no indication given as to the cut-out space to which it belongs, no mark can be given.

If more than one rectangle are indicated for the same cut-out space, no mark can be given even if one of the rectangles happens to be the correct one.

The actual markings used to indicate the solution offered are of no importance. Often the circle around the rectangle chosen is omitted and only the line drawn from the rectangle to the space where it is supposed to belong.

Sometimes numbers are used. Although we endeavour, in the presentation of the test, to assure uniformity in the method of response, any method clearly and without ambiguity indicating the subject's proposed solution is to be accepted if such solution is correct.

TEST NO. 4:

Two marks for every correct reproduction. The six small figures of Card No. 4 are regarded as a single sequence with a total possible credit of 2. In other cases where more

than one figure are contained on the same card, each receives two marks if correctly reproduced.

Possible Total: 24.

The following general principles apply to all the figures:

1. For the first mistake one mark is deducted. For the second mistake, the second mark falls away.
2. Neatness of drawing, firmness of line and general quality of the reproduction are disregarded.
3. Where a line supposed to meet another has been accidentally drawn too far no mark is deducted unless the overlap exceeds $\frac{1}{8}$ of an inch, when it is regarded as an error and a mark is deducted. The same applies when the line does not quite meet the other. A gap of $\frac{1}{8}$ " is allowed. If more, a mark is deducted.
4. Angles slightly rounded as a result of poor drawing may be accepted, but when clearly so intended a mark is to be deducted.
5. When a rectangle is required, a square will be allowed as a limiting case, but if the longer side becomes the shorter in the reproduction, a mark is deducted.
6. Any reversion of a figure is regarded as a mistake for which one mark is deducted.
7. In the case of any essential additions or omissions no marks are to be given.

Some types of drawings received are given below together with the marks allotted. Some cases will clearly be examples of the application of the rules given above, others are special cases. In each case a red arrow points at the mistake.

NO. I.

Full marks:



One Mark:



Rounded angle.




Proportion




Overlap more than $\frac{1}{8}$




Whenever more than one such mistake occur in the same figure, no marks are given.


No Marks:  Essential omission .


 Essential addition .



 Two mistakes

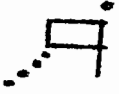

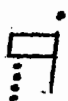
No. 2:

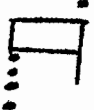
Full Marks:    One dot only below vertical line


One Mark:  One of the lower dots omitted.


 Lower dots increased to 5.

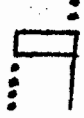
  Position of upper dot.

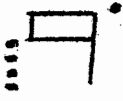
   Position of lower dots.

 More than one dot below the end of the vertical line.


 Vertical line extended below last dot for a distance more than twice that between two dots.

No marks:  Upper dot omitted.

 More than one dot above.

 More than one mistake.

No. 3:

Full marks: 

One mark:   The same mistake repeated symmetrically is regarded as only one mistake.





Proportion. Length of legs should be at least twice the vertical height of the feet.

No marks:



Essential addition.



Two mistakes.



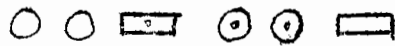
Essential omission.

No. 4:

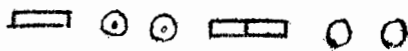
Full marks:



One mark:

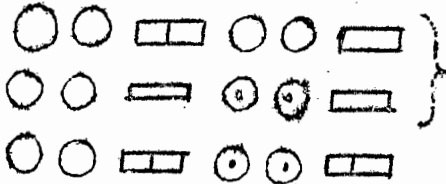


Proportion. Width of rectangle more than diameter of circles.



Perfect reversal

No marks:



Essential omission.

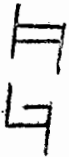
Essential addition.

NO. 5:

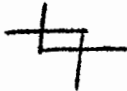
(First figure) Full marks:



One Mark:



No marks:



Two mistakes.



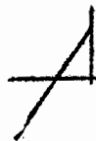
Essential Omission.



No. 5:

(Second figure)

Full marks:



One Mark:



The two portions a and b should be at least equal in length, otherwise b should be the longer.



Reversal.



Angle wrong.

No marks:



This is more serious than a mere gap. The essential nature of the figure is altered.



Two mistakes.

No. 6:

(First figure) Full marks:



One mark:



The diagonal nature of the shading has been retained.



One side more than 1 1/2 times another.

No marks:



The diagonal nature of the shaded portion is lost.

No. 6.

(Second figure) Full marks:



One mark:

If not noticeably of greater length than height.



...../No marks

No marks; If noticeably of greater height than length.

No. 7:
(First figure) Full marks:



One mark:



No marks:



No. 7:
(Second figure)

Full marks:



One mark:



Reversal.



No marks:



Essential additions.



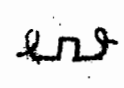
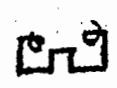
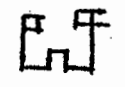
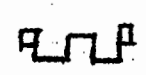
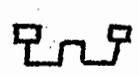
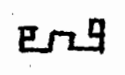
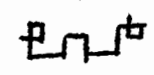
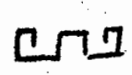
Essential Omission.



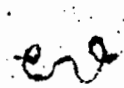
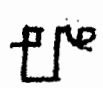
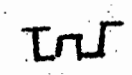
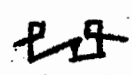
No. 8: (First figure).

(Following Terman) Full marks:

One mark:



No marks:



NO. 8:

(Second figure). We follow Terman's conditions for this figure. For full marks: (1) The outer figure must be a rectangle. (2) The inner rectangle must be off the centre to the right. (3) The inner figure may be a square but must not be noticeably higher than wide. (4) Lines from corner to corner must be drawn fairly accurately.

For one Mark: No essential part must be omitted or added. The design may be inverted, the inner rectangle in the centre or to the left; it may be taller than wide. Less accuracy will be accepted for the radiating lines, but they must show a tendency to meet the corners.

TEST NO. 5:

One mark for every identical pair indicated.

Possible total: 25.

(The last page does not form part of the test. No credit to be given for solutions on it.)

Indication of the two figures chosen may be of any kind: circles around the figures; crossing out of the figures; underlining; dots placed on them. Any method that leaves no doubt as to which figures are intended, is accepted.

Whenever more than two figures in the same row are marked and the markings allowed to stand, with no indication that some of them are supposed to have been deleted, no credit is given even if the correct pair is amongst those marked.

It sometimes happens that a definite system is followed, such as marking every first and last figure in each row, or always the first two.

This is a clear indication that no attempt was made to obtain any real solution. Correct solutions obtained in such a way are apparently accidental and not to be credited. Hence no marks are allowed for any page where such a system has been followed consistently, even though some items may be correct. (In some cases the first page shows evidence of real effort, but on the second a system - so much easier - is adopted.)

TEST NO. 6:

One mark for every correct solution.

Possible total: 25.

(The last page is not part of the test. No credit for solutions on last page).

For the correct solutions given below, the 5 drawings or figures in each row or item of the test are numbered from the left and the solutions given by numbers.

...../Correct

CORRECT SOLUTION OF TEST ITEMS OF TEST NO. 6.

ITEM (ROW) NO.	SOLUTION.	ITEM (ROW) NO.	SOLUTION.
1	4	14	3
2	1	15	1
3	3	16	2
4	5	17	2
5	4	18	5
6	1	19	2
7	4	20	1
8	3	21	3
9	4	22	5
10	3	23	5
11	5	24	3 4
12	2	25	5
13	3		

3.4. 1900

CORRECT SOLUTIONS OF ADDITIONAL ITEMS, NOT PART OF THE TEST.
NO CREDIT TO BE ALLOWED.

ITEM NO.	SOLUTION.	ITEM NO.	SOLUTION.	ITEM NO.	SOLUTION.
1	5	4	5	7	3
2	3	5	1	8	5
3	4	6	2	9	4

Whenever more than one figure or drawing is indicated in one row or item, credit is not given although one of the indications may be correct.

When a definite system is followed consistently, such as indicating all the first or all the last figures, no credit is given even for items solved correctly - purely by accident - in this way.

The totals on all tests are added in their original form to obtain the grand total. No weighting of the scores of the different tests has been introduced.

B I B L I O G R A P H Y.

1. Amoss, H.: Ontario School Ability Examination, Ryerson Press, Toronto, 1936.
2. Arthur, G.: A Point Scale of Performance Tests, I and II, Commonwealth Fund, New York, 1930.
3. Bieri, E.: Ein Beitrag zur Kenntnis der geistigen Entwicklung des taubstummen Schulkindes. Zürich & Leipzig, 1931.
4. Brown, A.W.: The Development and Standardization of the Chicago Non-verban Examination. J.Appl. Psychol., **XXIV**, 1940, 36 - 47 and 122-129.
5. Brown, W. and Thomson, G.: The Essentials of Mental Measurement, Cambridge, 1921.
6. Burchard and Myklebust: Comparison of Congenital and Adventitious Deafness, A.A. of the D., 1942, 140 - 154.
7. Burt, Cyril: The Factors of the Mind, Univ. of London Press, 1940.
8. Cattell, H.B.: A Guide to Mental Testing; Univ. of London Press, 1936.
9. Conference on the Problems of the Deaf, National Research Council, Washington, 1928.
10. Drever, J. and Collins, M.: Performance Tests of Intelligence Oliver and Boyd, Edinburgh, 1928; seconded, 1936.

11. Ewing I. and Stanton: A Study of Children with Defective Hearing, *Teacher of the Deaf*, 41, 1943, 56-59.
12. Fick, M.L.: An Individual Scale of General Intelligence for South Africa, Bureau of Educational Research, Pretoria.
13. Fisher, R.A. and Yates: *Statistical Tables for Biological, Agricultural and Medical Research*, Oliver and Boyd, Edinburgh.
14. Fisher, R.A.: *Statistical Methods for Research Workers*, Oliver and Boyd, Edinburgh, 1925 - 1948 .
15. Goodenough, F.: *The Measurement of Intelligence by Drawings*, World Book Co., New York, 1926.
16. Greenberger, H.: "Doubtful Cases," *Am. An. of the Deaf*, 34, 1889, 93.
17. Guilmartin, E.: "A Summary of Psychological Tests applied to the Deaf", The Volta Bureau, Washington, 1932.
18. Herderschêe: *Verstandsmetingen by Doofstomme Kinderen*, Maandschrift v. Verloskunde, Vrouwenziekte en Kindergeneeskunde, 1919, 569.
19. Höfler, R.: "Über die Bedeutung der Abstraktion für die geistige Entwicklung des taubstumme Kindes", *Zeitsch. für Kinderforschung*, XXXI, 1927.
20. Holzinger, K.: *Statistical Methods for Students in Education*, Ginn and Co., Boston, 1928.
21. Kilpatrick, W.M.: "Comparative Tests", *Am. Ann. of the Deaf*, 57, 1912, 427 - 8.

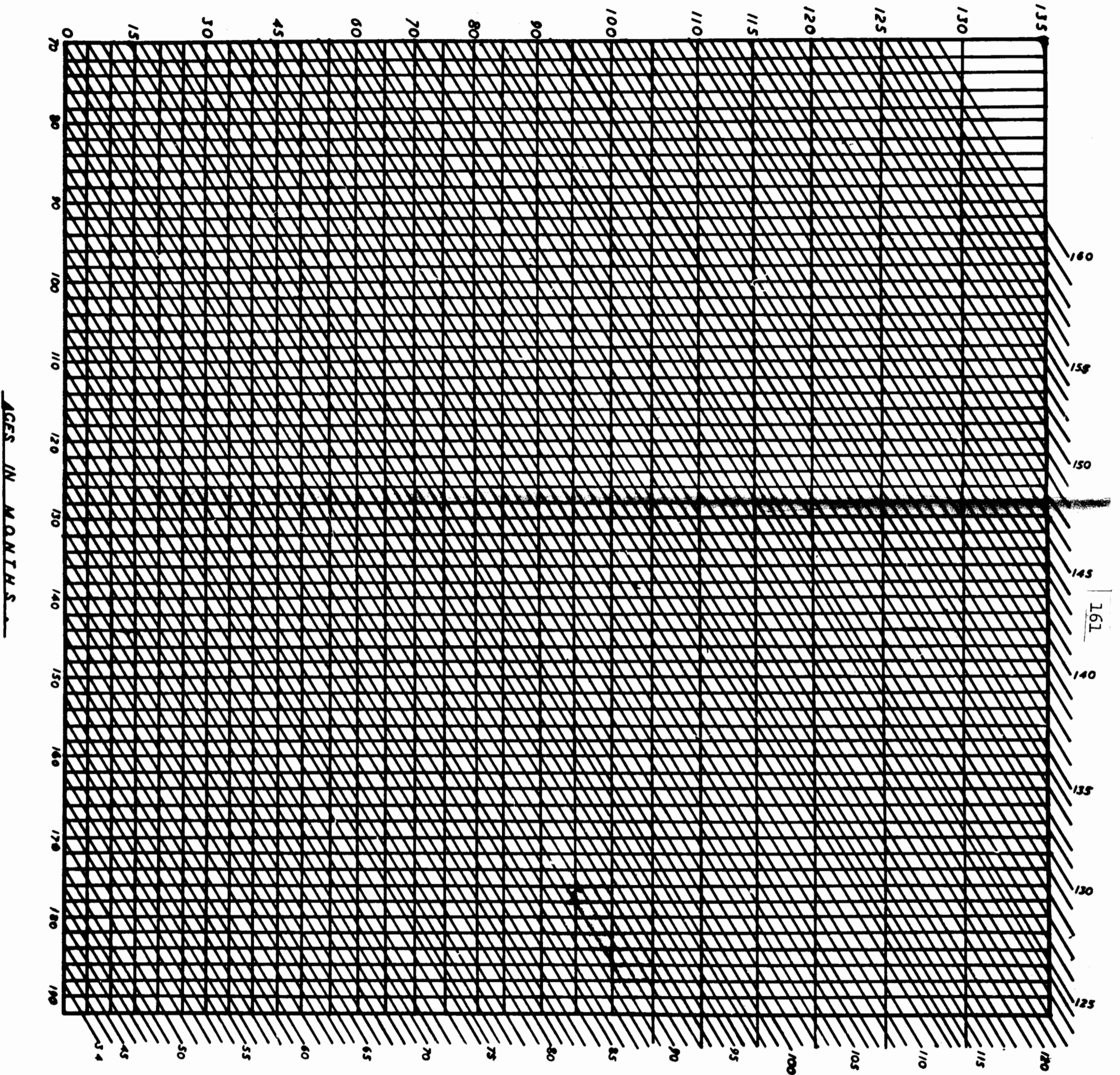
22. Kingsbury Primary Group Intelligence Scale, Public School Publishing Co.
23. Krenberger, S.: "Intelligenzprüfungen taubstumme Schüler", München, 1927.
24. Lane, Helen: "The Relation between Mental Test Scores and Future Achievement", Proceedings of the National Forum on Deafness and Speech Pathology, St Louis, 1947.
25. Lane, Helen: "The Education of the Deaf Child", in Twentieth Century Voice and Speech Correction, ed. E. Froeschels, Philosophical Library, New York, 1948.
26. Lindner, R.: "Vergleichende Intelligenzprüfungen," in Das Taubstumme Kind, Leipzig, 1925.
27. Lindquist, E.F.: Statistical Analysis in Educational Research, Houghton Mifflin Co., Boston, 1940.
28. McCall, W.: How to Measure in Education, McMillan, New York, 1922.
29. MacKane, K.: "A Comparison of the Intelligence of Deaf and Hearing Children," Teachers' College Contributions to Education, No. 585, Columbia University, New York, 1933.
30. MacMillan, D.P. and Bruner, F.G.: Special Report of the Department of Child Study and Pedagogic Investigation, Chicago Public Schools, 1906.
31. MacPherson, Jane: "A Comparison of Scores of Deaf and Hearing Children on the Hiskey Test", Washington University, St Louis, 1945.

32. Matz, W.: "Zeichen und Modellierversuch an Volksschülen, Hilfschülen, Taubstummen und Blinden", Zeitschrift für Angew. Psychol., X, 1915, 62 - 135.
33. Morrison, W.J.: "Ontario School Ability Examination", Am. Ann. of the Deaf, 1940, 184 - 189.
34. Mott, A.J.: "The Ninth Year of a Deaf Child's Life", Am. An. of the Deaf, 44, 1899, 401 - 412.
and 45, 1900, 33 - 39 and 93 - 109.
35. Oleron, P.: "A Study of the Intelligence of the Deaf", Am. An. of the Deaf, 95, 1950, 179 - 192.
36. Oliver, R.A.C.: "Mental Tests for Primitive Races", Year Book of Education 1935, Univ. of London Institute of Education, Evans Bros., London.
37. Oliver, R.A.C.: "General Intelligence Test for Africans", Education Department, Nairobi.
38. Pearson, K.: "On the Theory of Contingency and its Relation to Association and Normal Correlation", Draper's Research Memoirs, Biometric Series I, 1904.
39. Pintner, R. and Paterson, D.G.: "The Binet Scale and the Deaf Child". J. Educ. Psychol., VI, 1915, 201 - 210.
40. Pintner, R. and Paterson, D.G.: "Learning Tests with Deaf Children", Psychol. Monographs, XX, 1916, No. 88.
41. Pintner, R. and Paterson, D.G.: "A Scale of Performance Tests", Appleton, New York, 1917.

42. Pintner, R. et al.: "The Psychology of the Physically Handicapped", Crofts, New York, 1941.
43. Pintner, R.: "A Nonlanguage Group Intelligence Test", J.Appl. Psychol., 111, 1919, 199 - 214.
44. Pintner, R. Day and Fusfeld: "A Survey of American Schools for the Deaf", National Research Council, Washington, 1928.
45. Pintner, R.: "A Group Test suitable for Younger Deaf Children", J.Educ. Psychol., XX11, 1931, 360 - 363.
46. Porteus, S.D.: "The Measurement of the Intelligence of 653 Children Examined by the Binet and Porteus Tests", J.Educ. Psychol., IX, 1918, 13 - 31.
47. Reamer, J.C.: "Mental and Educational Measurements of the Deaf", Psychol. Monogr. XXIX, 1921, 132.
48. Smith, J.L.: "Mental Characteristics of Pupils", Am. An. of the Deaf, 48, 1908, 248 ff.
49. Snedecor, G.W.: "Statistical Methods", Collegiate Press, Ames, Iowa.
50. Snyders-Oomen, A.: "Intelligentieonderzoek van Doofstomme Kinderen", Berkhout, Nymegen, 1943.
51. Taylor, H.: "A Spelling List", Am. An. of the Deaf, 42, 1897, 364.
52. Ten Have, T.T.: "Over Psychologisch Onderzoek van het Dove Kind", Tydschrift voor Doofstommen-
onderwys, 1941.

53. Terman, L.M.: "The Measurement of Intelligence", Harrap, London, 1937.
54. Thurstone, L.L. and T.T.G.: "Factorial Studies of Intelligence", Psychometric Monographs No. 2, Univ. of Chicago, 1941.
55. Vertes, O.J.: "Memory in Deaf Mutes", Psychol. Stud., (Rauschburg - Festschrift), 1929, 210 - 216.
56. Wilcocks, R.W.: Die Suid-Afrikaanse Groep-Verstandstoets, Univ. v. Stellenbosch, 1931.
57. Wilcocks, R.W.: "On the Distribution and Growth of Intelligence", J. Gen. Psychol., VI, 1932, No. 2.
58. Wilcocks, R.W.: "The South African Group Test of Intelligence", Second Revised Edition, Pro Ecclesia, Stellenbosch, 1942.
59. Yoakum and Yerkes: "Mental Tests in the American Army", London, 1920.
60. Yule, G.U.: "An Introduction to the Theory of Statistics", 13th ed., London, 1948.
61. Zeckel, A., and Van der Kolk, J.J.: "A Comparative Intelligence Test of Groups of Children Born Deaf and of Good Hearing by Means of the Porteus Test", Am. An. of the Deaf, 84, 1939, 114 - 123.

ORIGINAL TEST SCORES.



1 Q LINES.