A STANDARDISATION OF PROFESSOR SPEARMAN'S " A MEASURE OF ' INTELLIGENCE ' FOR USE IN SCHOOLS " AND A DISCUSSION OF THE METHODS OF FACTORIAL ANALYSIS AS APPLIED TO THE STUDY OF ABILITY; BY J.E. BUTLER.

THESIS SUBMITTED IN PART FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF EDUCATION IN THE UNIVERSITY OF SOUTH AFRICA.

RHODES UNIVERSITY COLLEGE , 1948.

My thanks are due to the Principals and Staffs of the following schools for their co-operation in arranging for the testing , the results of which form the material for the experimental section of this investigation : Kingswood College , St Andrew's Preparatory School , St Andrew's College , St Aidan's College , Graeme College , The School of the Sacred Heart , The Diocesan School for Girls , Victoria Girls' High School , The Convent of the Assumption , The Grahamstown Primary School , St Joseph's Primary School and St Peter's Primary School.

I would also like to thank Professor E.H.Wild and Professor D.M.Morton , both of Rhodes University College , for the assistance and advice they gave me.

The numbers in brackets which appear from time to time in the text refer to the alphabetical list of references given in Appendix II.

INDEX

PART I

THEORETICAL WORK

1.Introduction.Soon after the death of Professor Spearman in 1945 P.B.Ballard wrote " And he worked to some purpose.In fact he gave us a new psychology - a new psychology of the intellect at any rate.It can scarcely be claimed that psychology is an exact science . It lacks both the definiteness of mathematics and the clarity of physics.It cannot be regarded - if indeed any science can be regarded - as a fixed body of knowledge capable of extension but not capable of modification.Its very foundations are vague and uncertain. We have , in fact , to view that curious entity the human mind through the eyes of a few men of genius. And in talking about it the very words we use were either invented by them or invested by them with new meanings to suit their purpose. Spearman was one of those master minds."(2). Page 1. This is no overstatement of the part played by Spearman in attempting to give to psychology the theoretical basis that it so badly needs and that clarity and accuracy which is characteristic of an exact science.Even though the theory of the physical sciences is still open to discussion and modification the physical scientist is able to predict with a high degree of accuracy and for those aspects of matter that he measures he has quantitative scales which have been found to be consistent and valid. Spearman attempted to make up for two deficiencies in the study of the mind viz., the lack of accuracy in the measurement of human ability and the lack of a system of explanatory principles to describe the working of the human mind.

He therefore postulated a"pregnant system of principles." (14) Page 29. Psychology has been working in reverse he said , a mass of experimental evidence has been collected and we need a set of qualitative and quantitative laws which will unite into one coherent whole all this data. The data which concerned Spearman most were those collected by scholars who have attempted to measure intelligence. The inaccuracy of these methods has very largely arisen from the fact that we have not defined clearly what it is that we are trying to measure.The hesitancy of the practising psychologist to make diagnoses on the results of testing has often been taken by the opponents of mental testing to imply that there is little value in such work . But there have always been difficulties in diagnosis and the more material available the greater will be the

cofidande with which the diagnosis is made.Therefore while mental testing does not give the complete picture it does give some information and the more concise it is made the greater will be its diagnostic value. The difficulties of the medical diagnostician , while great enough , are small compared to those facing the practising psychologist or vocational guidance officer. From a strictly utilitarian point of view therefore , the need for concise definition which will form an adequate basis for accurate measurement is an urgent one.

Spearman is also remembered for the part he played in the introduction of statistical methods into the study of psychological problems.He devised several new methods and his whole system of factorial analysis has stimulated an enormous amount of research. Many others in the same field have disagreed with his methods and conclusions but all pay tribute to the role he has played.He has , to use an analogy from electricity , played the part of an exciter by inducing a new field that has made possible the generation of a vast amount of new material. Thomson says " The main idea which still , rightly or wrongly , dominates factorial analysis was enunciated by him ..... and practically all that has been done since has either been inspired or provoked by his writings." (19)Page 5.

The basis of the theory was laid down by him in 1904 with the publication of " General Intelligence , Objectively Determined and Measured ". This was followed by a number of pamphlets and articles leading to the publication of " The Nature of Intelligence and the Principles of Cognition " and " The Abilities of Man " in 1923 and 1926 respectively.He collaborated with Holzinger , an American who has played an important part in the statistical analysis of educational and psychological data.

2.<u>Spearman's Theory of Cognition</u>.

As this theory is well known and has been discussed and criticised ever since it was first postulated a summary of it will be given here.

(a). Qualitative Aspect. He suggested three qualitative laws which he claimed can account for all these aspects of cognition.

1. Apprehension of Experience. Any lived experience tends to evoke immediately a knowledge of its characters and experiencer. Sensory stimulation leads to the production of mental states or sentience which in turn cause the mind to become aware of its own experience.

2. Eduction of Relations. The mentally presenting of two characters tends to evoke immediately a knowing of a relation between them. These characters which we may call fundaments and the relations educed between them may also act as fundaments and so lead to the eduction of relations between relations.

3. Eduction of Correlates. The mentally presenting of two any character together with any relation tends to evke immediately a knowing of the correlative character. This process, says Spearman, can be shown to be the basis of the understanding of language.

These three laws account for the ability of the mind to become aware of its own experience and having accomplished this to go on to create for itself new mental content by the eduction of relations and correlates. He defines the manifestations of these principles as firstly, noetic, in that they are self evident to the mind and secondly, noegenetic, in that in addition to being self evident to the mind they are able to generate new items in the cognitive field. All thought, says Spearman, proceeds according to these principles ; from any single experience there is no limit to the number of relations and correlates that can be educed from the characters of that experience. In Spearman's words : " Correlate eduction regulates the creation of all coherent writing speech, and even thought, whether verbal or wordless ..... from the beginning to the end everything intellectual can be reduced to some special casec of eduction , relative or correlative. " (14) Pages 284 and 300.

Spearman's " A Measure of ' Intelligence' " which is used in the second part of this investigation may be considered in the light of his qualitative principles. This is a verbal group test and is composed of seven sub-tests.

Test I. " Small - Little. If these two words mean the same or nearly the same thing write and 'S' as your answer ; if they are different write 'D'." Here two fundaments are given and the relation

between them must be cognised.

Test II."To strike means most nearly the same as what? To fool? To hit? To hurt? To break? " Here the process is one of correlate eduction . A relation of sameness is given and to an original fundament and the subject has to chose the correlate from the four alternatives given.

Test III."Ring. Stick. Wheel. Circle.See which word means something most unlike the rest."Here the subject may proceed to look for three words which can be linked by the same relation or he may begin with one word and proceed until he finds a word that bears to it the relation of difference. By seeing the relation of the word he has chosen with the others he will be able to decide which is the word which is most different from all the rest.

Test IV ."What was it made him eat? Thirst? Hunger? Pain? Joy? " Here we have correlate eduction from a complex fundament and relation.

Test V. " Stamps are put on what? Tables? Letters? Pictures? Trees? " Here a fundament and arelation are given and the correlate has to be educed.

Test VI. "White is to Black as Yes is to what? No? Perhaps? If? When?"Here a relation is presented in the first pair of words and must be educed. This relation is then taken and with the third fundament given the correlate is educed from the four alternatives given.

Test VII." If A is larger than B and B is larger than C , how does A stand as compared with C?Here two relations are given which act as fundaments and a relation has to be educed between them. The four alternatives given are all expressive of relations.

It will be seen that the most mental processess involved in this type of test can be analysed in terms of these principles. Many will be found to be a mixture of relation and correlate eduction.

(b) Quantitative Aspect. As an exact science is concerned with quantitative work no less than qualitative the following five principles are put forward to cover this aspect.Indeed , it is the quantitative work that figures so largely in physics that is responsible for its being called an exact science.

1.Mental Energy. Every mind tends to keep its total simultaneous output constant in quantity however varying in quality. The total output may thus be regarded as a product of extensity and intensity, , this product having a constant value. As intensity increases so will extensity decrease and vise versa.

2.Retentivity. The occurance of any cognitive event produces a tendency for it to occur afterwards. The manifestations of this law are of two sorts. i. Inertia.Cognitive events always begin and end more gradually than their apparent causes.

ii. Facilitation. Cognitive events by occuring ~~together~~ tend to re-occur more easily. An important corollary of this law is the law of association."Cognitive events by occuring together tend to re-occur with greater ease."

3. Fatigue."The occurence of any cognitive event produces a tendency opposed to it occuring afterwards."

4.Conative Control." The intensity of cognition can be controlled by conation."

5.Primordial Potencies ."Every manifestation of the preceding four quantitative principles is superposed on as its ultimate basis certain primordial but variable individual potencies. "

This , in the barest outline,is the theoretical background to Spearman's theory of intelligence. He summarises the whole process of cognition arising from a single sensory stimulation in the following seven steps:

i.Afferent nerves are excited.

ii.The impulses are transmitted to the sensorium.

iii.These impulses generate sentient states of mind.

iv.Cognition proper takes place with the apprehension of the sentience.(First qualitative principle.)

v.There is a gain in the intensity and determinateness of the apprehension.

vi.Eduction of relations and correlates. (Second and third qualitative principles.)

vii.The mental content is further increased by

s

supplementation.There is no limit to this last stage . Correlate and relation eduction, especially when aided by reproduction, lead to the building up of higher and higher levels of thought.

Having stated his theory of cognition we may now see how it applies to the theory of intelligence.A test of intelligence now becomes a measure of the ability to educe relations and correlates and tests must be designed to do this. Spearman points to the strange way in which the sample of tests for each age group varied in the original Binet test ; memory tests were included for some age groups but not for others. On what basis was such an arbritary sampling made he asks.This error he felt was due to lack of definition. No one had attempted to define the intelligent act.We now have a criterion on which to judge whether a test is likely to be a good test of intelligence or not. But this criterion,it must be admitted,is a qualitative one so that it will have little value in increasing the accuracy of measurement.It will however prevent the recurrence of the error alluded to above. A rigorous quantitative criterion for the assessment of the efficacy of a test has been suggested and will be referred to later. (See Page 9)

3.Spearman's Theory of Intelligence . This theory is best considered from two points of view, the statistical and/the psychological.

(i).Statistical.Thorndike and others found that various abilities as measured by tests tended to show consistently positive correlations between themselves and to explain this fact they introduced the notion of a general level of ability. To Spearman such a sampling of abilities is impossible as we have not made up our minds as to which abilities we are to measure.To explain these correlations Spearman postulated the existance of a general factor which takes part in all cognitive operations.The theory is best stated in terms of his famous tetrad equation : If a, b, p, q, represent four tests and r represents the correlation co-efficient obtained from correlating the performances of a hypothetical sample of individuals on any indicated pair of the tests then the following relation holds :-

$$r_{ap} \times r_{bq} - r_{aq} \times r_{bp} = 0 .$$

Whenever this relation holds, and only when it does so, throughout a table of correlations then every individual measurement of every ability can

be divided into two parts , "g" a general factor which is constant for any particular individual and "s" a specific factor which varies from one individual to another and from one ability to another. "g" has much greater relative influence in some abilities than in others and these relative influences are expressed as saturations or loadings.

The tetrad difference (the quantity on the right-hand side of the above equation is seldom zero in practice.This, says Spearman, is primarily due to sampling errors.By calculating the probable error of the tetrad difference and applying the criterion which demands that the difference must be at least five times the probable error before it may be regarded as significant it is decided whether the difference is due to sampling error or not.

The equation may be disturbed in another way. If in a battery two tests which are very similar are included the correlations will not show zero tetrad differences . This says Spearman is due to an overlap of specific factors because of the close similarity of the mental operations involved in the tests.Factors called group factors have also been isolated to account for this specific correlation.These occur in some but not all of a given set of abilities. Before a battery can be relied on to give an accurate measure of "g" it must be "purified "of these disturbers of the hierarchical order. The extent of the overlap between two tests is calculated by correlating the tests under suspicion with what may be called"reference abilities." The choice of these reference abilities need not be arbritary . All the tests in a battery besides those being examined for overlap are pooled giving two measures which are a combinations of several tests.

The efficiency of a test as a measure of "g" is shown by its saturation or loading which is always less than unity except in the case where we have a perfect test of "g" where the value would be exactly one.The significance of these saturations will be illustrated best by considering a fictitious table of correlation coefficients. This table represents the coefficients obtained from correlating six tests numbered from one to six .

g	g	1	2	3	4	5	6
g	1	r_{1g}	r_{2g}	r_{3g}	r_{4g}	r_{5g}	r_{6g}
1	r_{1g}		.72	.63	.54	.45	.36
2	r_{2g}	.72		.56	.48.	.40	.32
3	r_{3g}	.63	.56		.42	.35	.28
4	r_{4g}	.54	.48	.42		.30	.24
5	r_{5g}	.45	.40	.35	.30		.20
6	r_{6g}	.36	.32	.28	.24	.20	

This table shows what is known as hierarchical order; the coefficients in any column show a constant ratio throughout the column. In the diagonal cells "1" (unity) should be inserted as the correlation of a test must be, by definition, unity. But in these correlations the specific factors play a part and so they do not conform to the hierarchical order. If we imagine a test which is the perfect test , a test of pure "g", then its self correlation is also unity but it conforms to the hierarchical order. In the second column of the table will be seen this correlation and below the coefficients r_{1g} etc which represent the correlations of our perfect test with tests 1, 2, 3 , etc.; these correlations also run across the table in the second row. A "g" saturation is merely the extent to which the test under consideration would correlate with a test of pure "g". If we are able to calculate the values of r_{1g}, r_{2g}, r_{3g} etc., we would have the "g" saturations of these tests. Now as the table is hierarchical the following relation holds in the first two columns:-

$$\frac{r_{1g}}{1} = \frac{.72}{r_{2g}} = \frac{.63}{r_{3g}} = \frac{.54}{r_{4g}} = \frac{.45}{r_{5g}} = \frac{.36}{r_{6g}}$$

But a similar relation will hold between the second and the third columns thus:-

$$\frac{r_{2g}}{1} = \frac{.72}{r_{1g}} = \frac{.56}{r_{3g}} = \frac{.48}{r_{4g}} = \frac{.40}{r_{5g}} = \frac{.32}{r_{6g}}$$

From these two relations or with any other similar relations drawn from this table the values of the saturations can be calculated giving these results :- $r_{1g} = .9$, $r_{2g} = .8$, $r_{3g} = .7$, $r_{4g} = .6$, $r_{5g} = .5$, $r_{6g} = .4$.

It will be noticed that each correlation in the original table is a product of two of these saturations.

e.g. $.72 = .9 \times .8$.

$.30 = .6 \times .5$.

In general $r_{45} = r_{4g} \times r_{5g}$

The specific factors are calculated and must be of such a value that the sum of the squares of the saturations must be equal to unity. The six tests can therefore be described in terms of the following equations:-

$$z_1 = .9g + .436s_1$$
$$z_2 = .8g + .600s_2$$
$$z_3 = .7g + .714s_3$$
$$z_4 = .6g + .800s_4$$
$$z_5 = .5g + .866s_5$$
$$z_6 = .4g + .917s_6$$

Each z represents the score of some person in the test indicated in the subscript , a score made up of that person's "g" and "s" in the proportions indicated by the saturations. It will also be seen that the better the test the smaller is the role played by the specific factors. In the perfect test there is no specific factor at all.

When devising a battery it is possible to find one that as a whole will show a higher "g" saturation than that of any individual member. This is done by weighting each test in proportion to its "g" saturation using a quantity proportional to $\frac{r_{ig}}{1 - r_{ig}^2}$ where r_{ig} is the "g" saturation of a test "i". Several conditions must be observed. The tests added must conform to the hierarchical order already shown by the other tests. The specific factor in each test must apply to that test and that test only. In this way we are able by adding suitable tests to find a battery which is practically a test of pure "g".

We have now found the quantitative criterion for judging the efficacy of a test alluded to earlier. (See Page 6) Spearman has given us, for the first time, a means of judging a test that is based on more than mere opinion. His methods have been strongly criticised but he has drawn the attention of others to this problem and by the way in

which he has stated his theory he has challenged others to formulate other criteria which they have attempted to make equally unequivocal.

(ii) Psychological. The interpreatation of the results of factorial analysis has given rise to as wide a divergence of opinion as the factorial methods themselves.Theories of intelligence are a combination of certain psychological intuitions and experimental evidence. The limits of legitimate interpretation of the results of factorial analysis have too often been ignored .The evidence yielded by these methods is indirect and far less amenable to concise definition and description than much of the evidence on which physical and chemical hypotheses are postulated.The chief danger has been that we have tended to over-interpret , to reify the factors and give them isolated concrete existence.Spearman's theory lends itself to this tendency, a weakness for which he is not entirely blameless. His theory appears on the first reading of it to be an attractively simple explanation and for this very reason must be studied minutely.It must be remembered that the factors isolated have no concrete existence ; they are explanatory or descriptive concepts.They should not , on the basis of our present knowledge of them be equated with abilities.

The word "intelligence" has says Spearman been defined and re-defined so many times that it has lost all value as a scientific term.He therefore postulates ~~the existence of~~ a purely hypothetical factor which enters into all cognitive operations , this factor to be called "g".He interprets this factor by means of the concept of mental energy ,"g" representing the amount of energy that the mind has available for its cognitive operations. " The energy invoked to explain mental span is of an extremely hypothetical nature . As a concept it is incomparably less definite than that of the physical energy that serves the physiologist so well. And as for the relation between the two things , this is for the present buried in complete obscurity." (15)Page 407.This general factor varies in three dimensions from individual to individual.

(i).Each individual has a maximum quantity of energy available which is measured by "g".

(ii).Each mind has a certain degree of inertia peculiar to it.

(iii).Each mind has the faculty of recuperation after effortful expenditure.

The specific factors can be interpreted by making use of the analogy of "engines". Just as in the physical world engines are necessary to convert energy into work so in the mind there are engines through which the energy of the mind is converted into "work". These engines may be regarded on the physiological level as different neural systems.The successful solution of any problem confronting the mind depends on the potential of the energy and the efficiency of the engine concerned. Some operations depend more on the energy and others more on the engines for their successful solution.These relative roles are , of course , expressed in the factor saturations.

Spearman is hesitant to give any of the factors the name of abilities. He is content to regard them simply as factors.As regards the group factors already referred to (See Page 7) he finds that the number isolated is surprisingly small.The"job analysis " that has been applied in industry has led people to believe that a number of such abilities could be measured and defined.Five group factors have been found ,he says, that lead to specific correlations of significant magnitude.These are the logical , mechanical , musical , psychological and arithmetical abilities which he feels are definite enough to have important applications. There is a suggestion he claims that these factors are due more to past experience than native aptitude. This could be interpreted to mean that the engines can be made more efficient by training.

Spearman's methods and the interpretation of the results have been widely criticised .His claim for the universality of "g" has been strongly disputed and some scholars have held that no general factor exists. Coupled with the doubts expressed about the general factor have been the claims for the existence of a large numberof group factors.To some of these views we will now turn and show on what bases his theory has been criticised.

4.Thomson's Theory. All factorial work , says Thomson , rests ~~on~~ fundamentally on two experimental facts : Firstly most correlations between the performances of humans on mental tests ~~tend to be~~ are positive and secondly that square tables of correlations tend to be reducible to low rank when suitable diagonal elements are inserted.(When the rank of a table is one then the tetrad differences in that table are zero.) From the foregoing discussion it will be seen that the tetrad difference criterion is vital to the Spearman theory. Thomson , however , insists that far from being an established fact the zero tetrad difference is only a tendency . This tendency he holds has nothing to do with mind at all but is a mere mathematical necessity. If enough correlations are taken then the smaller tetrad differences will tend to swamp the larger and the equation will be satisfied. When the probable error criterion is applied to a tetrad difference and the difference is found to be negligible all that is proved is that it may be a sampling error from zero.A purified battery that shows ~~a~~ zero tetrad differences demonstrates nothing beyond the fact that it is possible to construct a battery that will show such differences and will thus be adequtely described in terms of two factors.The tendency to hierarchical order is not disputed; ~~but~~ the residues found in calculating " g" saturations have not been tested to see if there is not a probability of their being not due to sampling error.

Another objection based also on the sampling theory is ~~stated~~by him thus " (Sampling errors) make it difficult for us to see the true outline and easy to entertain hypotheses that cannot be disproved , though often they cannot be proved either by the data." (19)Page 181. When we standardise a test we test only a sample of the population and the decision as to whether the sample we have chosen is a representative one is made on purely conventional grounds based on the theory of probability.It seems therefore that the factors we obtain from test results will depend on the sub-population we have tested. After considering the effects of both univariant and multivariant selection he concludes "All these considerations make it extremely doubtful indeed whether any factors have absolute meaning . They seem to be entirely dependent on the population in which they are measured , and for their definition

there would be required not only a given set of tests and a given technical procedure in analysis but also a given population.(19) Page 194.

On the whole Thomson uses the same statistical methods as suggested by Spearman but he is not prepared to limit the number of the general factors in the same way.He believes that group factors can be used to advantage and claims that Spearman by minimising the number of general factors hasmaximised the influence of the specific factors. He claims that the appearance of group factors (others as we shall see later have accepted them) is fatal to the two factor theory. All the evidence yielded by this analysis can be explained in terms of a sampling hypothesis.The mind should be regarded as a homogeneous , structureless organ consisting of a large number of "bonds".Each test calls on a sample of these bonds ; some of them are common to two or more tests and cause their correlation.Tests are assigned coefficients of "richness" which are a measureof the fraction of the total number of bonds in the mind concerned in the particular operation.

Hierarchical order on this theory is the most probable order but not the only order.If tetrad differences were zero then we would have to accept "g".This theory is able on the other hand to explain the large terad differences found in unpurified "batteries; the sampling of a number of bonds common to two tests would result in a higher correlation between these and would thus disturb the tetrad difference. These high correlations do not mean that the whole mind is being called on but that we are using " sub-pools" of the mind.These can be used also to explain the common factors of the Thurstone theory. (See Page 16)

This theory does not , as suggested by Spearman, make all men equal but merely that each mind contains a sample of a possible number of bonds"... each man possessing some , but not all , both of the inherited and the aquired neural bonds that are the physical side of thought." (19) Page 54. Divergences from hierarchical order indicate structure ; experimental work has shown that the number of factors shown by adults is greater than that shown by children which seems to suggest that training has imposed on the mind a structure which is absent in youth.There is an absence of fixed linkages in the mind so that any sample whatever of the bonds can be assembled in the

activity called for by the test. " It is not without significance that the factor most widely recognised after Spearman's "g" is the verbal factor "v" , the mother tongue being , as it were, the physical body of the mind , its aquired structure. " (19). Page 313.

Under certain conditions it is possible to have a battery of tests which has a "g" saturation ofgreater than unity. This is known as a "Heywood Case". Such a saturation is by definition impossible on the two factor theory.If such a case is found that conforms to the hierarchy then the hierarchy must be abandoned as the basis of the theory. This narrow dividing line between the perfect and the impossible makes the psychological interpretation of our mathematically derived information extremely difficult.

In conclusion , " The author's (Thomson's) attitude is that he does not believe in factors if any degree of real existence is attributed to them ; but that , of course , he recognises that any set of correlated human abilities can always be described by a number of variable s or factors and that in many ways (Thomson's italics), among which no doubt some will be more useful or more elegant or more sparing of unnecessary hypotheses .But the mind is very much more complex , and also very much more an integrated whole , than any one naive interpretation of any one mathematical analysis might lead the reader to suppose." (19) Page 299. Thomson's theory is a multiple factor one ; he regards the proof of the existence of "g" as insufficient and prefers to describe correlations in terms of a number of group factors.Finally,as stated above,the danger of reifying the factors is agreat one , they have no more existence than a cost-of-living index.

5.Thurstone's Theory. Thurstone felt , as did others , that Spearman's account of specific correlation could not be accepted.The weakness of the explanation lies in the fact that when a significant tetrad difference is found the factorist tends to look for undue similarity in the tests which cause this overlap which he was not able to see in the first instance. Spearman gives us no basis for the judgement of this similarity nor does he attempt to explain ~~the~~ it in terms of his qualitative principles and it is difficult to see how this would be done .The notion of overlap as an explanation of divergences from zero tetrad difference does not appear to be acceptable .As the tetrad difference criterion is essential to the Spearman theory it was essential that an explanation be found for any divergences from it.Thurstone considers the Spearman two factor theory ~~as only~~ valid only under certain conditions and thus unable to account for all cases.

The basic theorem of the Thurstone theory may be stated thus : The correlations derived from a number of tests may be described by as many common factors as the reduced rank of their correlation matrix. A matrix has a reduced rank of one when the tetrad differences are zero , a reduced rank of two when the tetrad difference of the tetrad differences is zero and so on. When the rank of a matrix has been established (the methods of doing it do not concern us here) values are calculated for the diagonal cells which are of such ~~a value~~ an order as to partake of the same rank as the rest of the matrix.These values are inserted and the columns summed.All the sums are added together and the square root of the result found.The total of each column is then divided by this root giving results which are the saturations (loadings) of the first common factor.These factors are then written vertcally and horizontally and a new matrix calculated by finding all the ~~sums~~ products.These values are subtracted from the ~~origina~~ corresponding members of the original matrix ~~and~~, the residuals are then treated as ~~though~~ it they were a table of correlations and the second common factor is found in the same way. This process is repeated until the residuals disappear of are insignificant. Once thee common factors have been found the specific factors are found in the samecway as done in the Spearman method: the sum of thw squares

of all the factors on any test must be equal to unity.

We get from this method a number of common factors and a specific factor for each test. In a nonad where the tetrad differences of the tetrad differences is zero the matrix is deecribed in terms of two common factors and a specific factor for each test.In a tetrad which has a zero difference the conclusion is identical with that arrived at on the Spearman theory viz., one common factor and a specific factor fof each test. If the first common factors are written down and the products found the original matrix will be givenvand so the residuals are zero. One common factor is therefore sufficient in this case. In cases where there are several common factors some of the factors will be negative in almost every case.

Thurstone's method leads ideally to what he calls "orthogonal simple structure" whereby he means that his factors give a unique description of the correlations and these factors are uncorrelated and distinct.He insists that before the factors can be given any psychological significanse all negative values must be removed.This is accomplished by a process known as "rotation.Translating the correlations into geometrical terms he conceives of them as vectors which when they are atvright angles are uncorrelated.With n factors he thinks of an n - dimensional space in which the common factors if they are uncorrelated are at right angles to each other. Within this space by kepping the vectors at right angles a rotation is made until there are no negative values at all.We thus arrive at set of factors that are all positive and uncorrelated ; orthogonal simple structure.But there are cases where this structure is not possible.and to account for these cases he introduces"oblique"factors which are correlated to a certain extent. With these he says it is still possible to arrive at a simple structure , a unique definition even if the factors are slightly correlated. Certain conditions must be observed in the doing the rotation ; no test may contain all the common factors , no common factor may occur in all the ~~common-factors~~ tests and the loadings must be in such a pattern that all the test are qualitatively different from one another.

The results of factorial analysis may be used in a number of ways.A theory can be based on the results~~/or the test can be used to~~

or the results can be used to test a theory.Thurstone has from the beginning regarded the mind as analysable into a set of distinct abilities that are responsible for an individual's performance on a particular test.

" Factor analysis is reminiscent of faculty psychology.It is true that the object of factor analysis is to discover the mental faculties .The severe restrictions that are imposed by the logic of factor analysis make it an arduous task to isolate each new mental faculty, because it is necessary to prove that it is called for by the experimental observations." (23).Page 53.He is thus committing himself to a theory of mind which has been criticised very strongly by others , and he is assuming that the factors he isolates may be equated with the abilities he believs function in the mind.For this reason he insists that the results of factorial analysis must be psychologically meaningful." The identification of factors cannot be accepted unless the factors make psychological sense ..... Factorial analysis aids in discovering the principal dimensions of the experimental observations , and it is our further task to discover the psychological meaning of each factor. Only to the extent that we succeed in isolating meaningful categories are we advancing psychological science." (22) Page 90.

He does not accept Thomson's analysis of the effects of selection and hopes to be able to show that the factors do not change even if the test is moved to another battery and also that these factors will remain constant from one population to another.~~For~~Because the principal components used by Hotelling do not show this invariance he has discarded that type of solution.The Hotelling factors change when a tsst is added to or removed from a battery. "The solution I have found for this problem is to rotate the reference frame into a simple structure if it exists in the battery.To increase the number of verbal tests or to with draw some of the spatial tests , for example , has no effect on the reference frame for a simple structure.This type of invariance gives the possibility of identifying the reference axes in terms of the unique traits of the individual subjects." (22).Page 91.

In 1937 Thurstone conducted an investgation with the object of isolating the primary mental abilities that he believes can be described by factorial analysis. Two hundred and forty subjects

were tested with a battery of fifty six tests which were assembled so as to represent a fairly wide range of mental activities that are typical of current psychological tests , with special emphasis on those tests which are used as measures of general intelligence " (22) Page 910.The results of the testing were analysed and the following factors isolated and defined : A spatial factor S , a perceptual factor P, a number factor N , two verbal factors V (ability to grasp ideas and meanings) and W (fluency in dealing with words) , a memory factor M , an inductive factor I and two factors D and R which Thurstone is not able to identify at this stage. All these factors will , he believes , be of great value in vocational guidance when our methods of measurement and interpretation are made more accurate. " Further experimental study of the psychological nature of these factors ~~is necessary~~ together with other information about the subjects such as their likes and dislikes , their occupational preferences and their actual performances will eventually give us more confidence in the interpretation of their mental profiles." (22) Page 102.

This theory is then a multiple factor one and is able to explain divergences from hierarchical order in a more satisfactory way than the Spearman theory.Thurstone hoped originally that orthogonal simple structure would be possible in all cases but he found that factors were often correlated. In these oblique factors Thomson sees the possibility of reconciliation between the Spearman and Thurstone points of view .When these factors appear it is possible that there is a general factor causing the correlation.By factorising the factors and finding second order factors we might find a general factor.If this is so Thurstone will have to admit a wider occurence of the general factor and the Spearman group will have to admit the universality of "g" as not proved.

6.Burt'sTheory. Burt has made his contribution in this field in the same distinguished manner as he has done in other fields of psychology.In 1917 he suggested the centroid method of calculating saturation coefficients that was later adopted by Thurstone .Later he suggested another method which he claims is superior to the others.This method is based on the least squares method of curve fitting.where the best fit is obtained when the sum of the squares of the residuals has a minimum value.The product moment formula for finding correlation coefficients yields a least squares fit to the two variablesbeing correlated.If we use the product moment method with a table of correlation we will thus get a least squares fit i.e.our factors will give the best representation of the table.If we square the original scores (in standard units) we can deduce the correlations and by squaring the correlations we can deduce the saturation coefficients.We therfore square the whole table of correlations and sum the squares.For a test j the saturation r_{gj} is calculated by summing all the squares in the j column , taking the square root and dividing it by the fourth root of the sum of all the squares of the correlations in the whole table.This method is,he claims,economical and accurate.

He criticises the Thurstone methods as inaccurate and laborious.Accurate results will not be obtained,he says,by artificially reducing therank of the matrix but by maximising the influence of all the factors and rejecting those that are statistically insignificant.The process of rotation as suggested by Thurstone is complicated and loses in accuracy as it proceeds .Burt can see no reason why the negative saturations should not be accepted. " Now I venture to suggest that this postulate - that intelligible factors or loadings components can have nothing but positive loadings or saturation coefficients - comes dangerously near to an illegitimate identification of factors with faculties..... The assumption is that the empirical correlations between the tests intended (Burt's italics) to measure'abilities' can beanalysed into abilities without remainder." (5).Page 57.

Muchdifference of opinion has been due to the fact that the problem has been tackled from different points of view even though identity of purose is acknowledged.Spearman is concerned with a theoretical

analysis of the mind while Thurstone and Thomson are more concerned with the prediction of individual progress and development.The primary purpose says Burt is exact and systematic description.To add to the confusion created by the different schools there have been cases where generalisations have been made from the results of a few correlation tables.In psychology it is seldom known when we are putting an hypothesis to the test whether we have excluded all the influences except those with which the hypothesis is concerned.In the light of this uncertainty the overzealous manner in which unique solutions are sought for complex problems appears to lead to an attitude of mind that makes the discarding of invalid theory difficult.The rejection and formulation of hypotheses is essential to science but we are apt too often , because we are unable to suggest an alternative explanation,to regard our own view as the only one possible.

Burt also warns against the reification of factors and holds that in our present knowledge of them we cannot assume a causal connection between them and the mind."My own view is that what holds of the correlations also holds of the underlying factors ; its discovery might reinforce a pre-existing presumption in favour of a causal connection but it cannot by itself create such a presumption.....The logical presuppositions of factor analysis afford no prima facie ground for treating the resulting factors as causal entities." (4). Pages 68 and71.

A theory of four factors is the fundamental logical postulate from which all factor analysis must start he says. These are as follows :

A. Common Factors.

1.Universal or general factors.

a.Positive or one signed factors.These have only positive saturations.Usually there is only one and is called the general factor.

b.Bi-polar or two-signed factors .These may have positive or negative saturations.

2.Particular factors. These are common to a group of factors and are thus called group factors.They have been called overlapping specific factors but this merely blurs the meaning.

B. Individual or unique factors.which influence one test or trait alone.

3.Singular factors peculiar to a single trait and called specifics.These may be constant or systematic errors.

4.Accidental factors. Eachnof these is peculiar to the particular occasion on which the particular trait is measured.

"The measurement of any individual for any one of a given set of traits may be regarded as a function of four kinds of components , namely those characteristic of :

i. All the traits;

ii.Some of the traits ;

iii.The particular trait in question whenever it is measured .

iv.The particular trait in question as it is measured on this particular occasion." (4) Pages 102 & 3.

All other theories , says Burt , are merely simplifications of this theory and by the very nature of simplifications are unable to account for all the observed facts.All factors must be regarded as group ones with the general and specific factors as extreme and limiting cases.

As stated above he condemns the equating of factors and abilities. The problem of definition is a considerable one and he is extremely cautious in dealing with this problem.To disprove the concrete existence of a factor is not to abolish its value ; it is a term of reference only . As it is difficult if not impossible to define accurately the traits we mea wish to measure some attempt at measurement is necessary for the accurate specification of the traits." The introduction of quantitative concepts intompsychology is " merely an attempt to make our efforts at systematic classification more and more precise." (4) Page 138.Hecdescribes the factors as follows : "So far as it seeks to be strictly scientific , psychology must beware of supposing that these principles of classification can forthwith be treated as factors in the mind e.g. as primary abilities or as mental powers or energies.Factors specify not unitary qualities but systematic patterns ; not active entities , but relations between what we loosely call the mind and

what we vaguely call its environment; i.e.they specify systems of relations between two sets of relational systems." (4) Page 251.

Burt accepts "g" and would give it and other factors as much existence as the concept of physical energy can claim.He defines "g" as "innate general cognitive efficiency.....as specifying certain individual differences in the structure of the central nervous system" (4) Page 216. He lays greater stress on the organisation of the mind rather than on a number of bonds or an amount of energy.Minds differ in their inborn capacity for complex organisation in the tissue as yet unorganised. Stated differently by Garnett but accepted by Burt " 'g' measures my power to concentrate my-attention when I wish to do so ." (4) Page 217.

Burt regards the specific factors as being constant or systematic errors.They show an extremely low stability from one investigation to another. Spearman would agree with this as he regards "g" as the main factor in intelligent behaviour; but goes further in saying that the measurement of specific abilitiesfor the purposes of vocational guidance would be like the blind leading the blind. Burt disagrees on the basis of experience ; and holds that group factors can be extremely useful. In later years Spearman came to a substantial acceptance of Burt's view.Spearman in the first statement of his theory admitted the existence of a few group factors but felt that the number should be kept at a minimum.

Burt's is thus a four factor theory. He accepts "g" but regards the group factors as not only existent but most useful. He has shown on mathematical grounds that many methods that appear to be different do in fact yield sustantially the same results.His theory might form the basis of a reconciliation of the various theories.

7.Holzinger's Bi - factor Theory. Only a summary of this theory will be given here.Holzinger admits the existence of "g" running through all the tests and attempts to account for the remaining variance by using a set of mutually exclusive group factors.It is thus a natural extension of Spearman's theary as it embodies his fundamental distinction between general and non-general factors.Saturations with "g" are calculated from batteries that form perfect hierarchies , the correlation due to "g" is removed th and the residuals treated to find the group factors.

Adiagram of such an analysis is given below ."g" runs through all the tests and the factors h,k,and l run through small and mutually exclusive groups

	Factors			
Tests	g	h	k	l
1	x	x		
2	x	x		
3	x		x	
4	x		x	
5	x			x
6	x			x

Holzinger collaborated with Spearman in several publications dealing with the statistical problems raised by this analysis.Recently he has come around to Burt's view that the group factors play the dominant role.

8.Hotelling's Method of Principal Components. This method ~~has been~~ was alluded to in the discussion of Thurstone's theory .(Page 17.) If the scores ~~on two tests~~ of a sample population on two tests are used as co - ordinates and these plotted on two axes at right angles to each other the shape of ~~the~~ the area covered by the dots representing the scores of each individual on the two tests will be roughly elliptical if the two tests are correlated.This ellipse has two axes , a major and a minor.Similarly if there are n tests we would have an n -dimensional ellipsoid that has n principal axes.The principal components are found by calculating the axes along each principal axis giving a set of factors that diminish in influence as the number of factors increases. The first factor or component runs along the greatest length of the crowd and so is best able to indicate an individual's position in the crowd. This method has one great advantage in that it is possible to reproduce all the original test scores if the factors are given.but it is not possible to reproduce the correlations shown with other tests as can be done with factors found by the centroid method or the method of least squares. It has one great disadvantage in the fact that as soon as the composition of a battery is changed all the factors change as well.Burt regards these factors as superior to those yielded by the Thurstone method and claims that the first three factors are able to account for a far greater amount of variance than the first three Thurstone factors.They are also truly independent whereas the Thurstone factors are always correlated to a certain extent.Thurstone,as would be expected,has criticised these factors as he feels they are psychologically meaningless.

9 Summary.Professor Spearman postulated a theory of cognition and derived from it a theory of intelligence.To cover the qualitative aspects of c cognition he suggested the following three principles.

i.Apprehension of experience. (Page 3).

ii.Eduction of Relations (Page 3).

iii.Eduction of Correlates (Page 3).

To cover the quantitative aspects he suggest five principles.

i.Mental Energy. (Page 5).

ii.Retentivity. (Page 5).

iii.Fatigue . (Page 5).

iv.Conative Control. (Page 5).

v. Primordial Potency. (Page 5).

Intelligence tests should be constructed against this theoretical background .Tests may be judged and compared by calculating their "g" saturations .

Dealing morecspecifically with his theory of intelligence we considered it from two points of views statistical and psychological. (Pages 6 and 10 respectively.)Statistically the theory rests on the er criterion of the tetrad difference .This difference shows a tendency to zero a fact that is shown by experimental evidence but whose interpretation is subject to discussion.When the tetrad difference is zero or insignificant Spearman claims that each individual performance may be described in terms of two factors : one general that enters into all cognitive pperations and one specific to each test.In psychological terms he describes the general factor by using a concept of mental energy and the specific factors by engines in the mind which convert the mental energy into "work". These terms are however purely explanatory no degree of real existence is postulated.Pages 10and 11.

A short summary of the methods used in calculating the values of general and specific factors has been made. Page 6 - 10.

The views of Thomson (Page 12); Thurstone (Page 15), Burt (Page 19) , Holzinger (Page 23) and Hotelling (Page 24) have been given in summarised form.

10.Conclusions and Criticisms. It is difficult to over-estimate the service Spearman has done in attempting to give psychology a sound theoretical basis which would make it less of a curious hybrid of science and art.Experiment without adequate definition was to him the cardinal weakness in a subject where the difficulties of measurement are great enough without being added to by a lack of clear objectives.He has drawn largely on the findings of philosophers and by ~~combing the~~ combining this knowledge with the findings of the experimental psychologist has attempted to formulate a theory which has an adequate basis in both spheres.In every study it is essential,that some master mind attempts to bring to bring together the results of many scholars who while concentrating on specific topics tend to lose sight of the relation that their work bears to the work of others on the same subject but on different aspects of it.It matters not whether the ~~encye~~ encyclopaedic view suggested by the master mind is acceptable or not; the very fact that he attempts to fit the results ofoothers into a coherent and logical theoretical framework causes the others to re - examine the results of their work and possibly to suggest an alternative explanation.In the field of mental testing Spearman put forward such a view and by so doing stimulated research to a remarkable degree.

The most stimulating of the views he put forward ~~was~~ is his method of analysing the relations between correlations derived from mental testing.Factorial analysis was first suggested by him and thus , quite naturally , the greater part of the earlier work was done in England.Later scholars in America entered the field and the volume of work done on the subject has ~~increased~~ expanded rapidly .This method ~~res~~ is based on two experimental facts : I.Correlations between the performances of a given group on many different tests tend to be positive.~~and~~

II.The tetrad differences of these correlations show a tendency to zero.

This ~~criterion~~ tetrad difference was taken by Spearman as the basis of his theory of two factors and the validity of taking this basis has been seriously questioned.Thomson has shown very clearly that the zero tetrad difference is what we would expect and claims that

only a tendency to zero is shown. Spearman's attitude to this difference was far more definite than this ; his theory of the general factor was based on those combinations of tests that showed zero tetrad difference. It would appear therefore that his acceptance of this criterion as the basis of a theory of intelligence is open to question.

Thomson after considering the effects of selection feels that the effects of these influences are great enough to doubt whether any/factorial structure exists. Burt agrees that selection has a considerable influence but suggests that this has been exaggerated and that while the factors may change in value they do possess a definite structure. Thurstone is also confident that an invariant structure exists. All these considerations make it clear how tentative must be the conclusions that we draw from these analyses .

There is a danger that we will build an inverted pyramid of calculations from a few data leading to a top-heavy structure which the original data are not able to support. It may be that the tests of significance that have been used have not been used rigorously enough or stand in need of revision.

Every statistical text when dealing with correlation warns the student against attributing a causal connection between two variables that are correlated. Factors are often isolated after elaborate treatment , the assumption made that these factors are caused by the mind and a theory sought for that will account for these "factors of the mind." As these factors are derived from correlations the same limitations of interpretation must apply to them.

Spearman's original insistence on the universality of "g", a view that he later modified , has not been accepted by later investigators. The view accepted to-day by a large group of factorists is that there may or may not be a general factor , that there are group factors which are the most important and that there are may be specific factors which apply to each type of test. Spearman felt that the group factors had to prove their right to existence and this they appear to have done.

The notion of overlap of specific factors to explain the divergence from the zero tetrad difference has been questioned by many. Thomson feels that there is a real danger that the factorist

will look for similarity after the ev factors have been isolated. On what basis is such a judgement to be made ? No one has drawn up a scale of similarity if such a thing were possible. We are , therefore , if we accept this explanation , introducing into a quantitative study a qualitative judgement whose basis is ill-defined.

The specific factors are regarded by many as errors ; they are , when the Spearman method is used , what is left over after the calculation of the general factor.Burt questions the validity of regarding these factors as capable of supporting the analogy of the engines.Spearman insists that these factors must not be called specific abilities . Elsewhere , however , he states that at the core of each there may be what may be described as " pure innate specific aptitude". It would appear therefore that while he is alive to the danger of equating factors and abilities he is really hankering after some such description of them.He is therefore not entirely blameless for the over - interpretation of the theory. His theory is attractively simple . We must beware of accepting a simple explanation in a subject where it is likely that simple explanations and theories will not be feasible at any stage.

The confusion that resulted after the first criticisms of the theory was due largely to a lack of understanding as to the nature of the evidence.The evidence yielded is indirect and the units used relative not absolute.While it may indicate lines of thought of a psychological kind it does not necessarily do so. The limits of interpretation are now more fully understood and factor analysis is now regarded as a distinct mathematical method to which the study of intelligence has no prior claim.Burt says " Let us agree then that mathematical analysis by itself cannot ever disclose what factors make up the human mind or what mental types are discoverable in the human race." (4) Page ~~99~~209. " Finally it is argued - instead of the psychologist invoking principles and postulates appropriate only to the simpler sciences the more complex sciences might in their turn profitably borrow factorial methods from psychology." (4) Page251. Burt has used these methods in the study ofvtemperament.

Burt has suggested a theory that is able to account for

all the factors that have been isolated.By regarding the factors as en entities distinct from the mind and thus capable of classification on a general basis he has at the outset given the warning that they are descriptive terms and that factors can be accepted even if negative as no psychological significance is inherent in them.Factor analysis is thus a method for finding the relations between correalations and the more diverse the methods we use the greater will be our knowledge of these relations.Spearman in suggesting the method of factorial analysis created a method of which can be used in a far wider sense than he realised and whose results we admit of a greater variety of interpretations. than he would,at first,admit. His complete and stimulating statement formed, in spite of limitations, the starting point from which all later work has proceeded.

Finally,it must be realised that each method of factorial analysis has its own merits and may be the best that can be used for a particular problem. " It is a positive fact that every one of the types of factorial solutions in extensive use to-day is an effiecient one.When known thoroughly , a skilled analyst can apply any one of them to produce fine results in a desired situation . The enthusiastic partisan who depreciates the efficacy of some other method and calls it a poor one thereby furnishes indisputable proof of one of two things - either his own sincerity or his own ignorance , not just ignorance of the details of the other method,but general ignorance. Any person of intelligence should know that it is highly improbable that any of the brilliant psychologists and statisticians who have worked on these problems for the last thirty years would produce and advocate a "rotten" system." (12) Page 8.

PART II

EXPERIMENTAL WORK

1.Introduction. The object of this investigation is the establishment of norms for Spearman's " AMeasure of 'Iintelligence' For Use in Schools".These norms are based on the results of testing all the English speaking children at in the classes from Standard III to Standard VIII inclusive or the equivalents of these classes , who were at school in Grahamstown in April 1947. The test , a verbal group one , has been discussed in the theoretical section of this work (See pages 3 and 4) . It is made up of seven sub - tests divided as shown below in Table I.

Table I. Sub - divisions of the Test .

Sub-test.	Type of Question	No of Questions
1.Same or Opposite.	"Are small and little nearly the same or about as different as possible ?	30
2.Synonyms	"To strike means most nearly the same as what? To fool? To hit? To hurt? To break?	24
3.Classification.	"Ring .Stick.Wheel. Circle. Which word means means something most unlike the rest?	28
4.Questions.	"What was it made him eat? Thirst? Hunger? Pain? Joy?	12
5. Completion.	"Stamps are put on Tables? Letters? Pictures? Trees?	24
6.Analogies.	"White is to Black as Yes is to what? No? Perhaps? If? When?	28
7.Inference.	"If A is larger than B and B is larger than C how does A stand as compared withC ? Smaller? Larger? Better? Equal?	17
	Total	163.

At the beginning of each sub test (except sub test 7)where there is only one) are two questions which are worked out with the class as a demonstration of the type of question in that section.These questions are included in the totals ; the actual number of questions

actually answered is 150 not 163.The instructions given with the test include a "patter" which must be used to introduce eack test so that we can be sure of standardised pr cedure.which is essential if we are to keep the response errors at a minimum,

(iii). Procedure. Several modificat..ons of procedure were introduced.

i.The instructions for tests 1 and 2 state explicitly that each question should be giventw- twice but in the other tests no mention is made of the number of times the questions should be given. It seems obvious that Spearman intended that all the questions be given twice but it was decided to interpret the instrictions to mean that questions in the/first two tests were to be given twice each and the rest of the questions once each. It was decided to apply the test in two school periods of forty minutes each so that the testing would involve a minimum of dislocation of the time table.To judge how long the test would take to apply and to familiarise myself with it I applied the test to a class of coloured children at the Rhodes University College Practising School. It was found that if each ~~test~~ question were given twice the test would take about ninety minutes tp apply. As this period was too long to make the testing possible in two school periods it was decided to shorten the time of application by applying the test as stated above.with the result that the test was applied easily and conveniently in two periods of forty minutes each.

ii.Spearman's instructions suggest the use of foolscap ; which would require folding and numbering by the subjects . As valuable time would be lost in this way forms were printed in such a way that subtests 1 to 4 were on one side of the sheet and the rest of the test on the other.Forms were taken in at the end of the first period of testing i.e,after the completion of sub-test 4 and were handed back in the second period,No reference was allowed,during the second period of testing , to work done in the first period and while it is possible that mistakes made in the first period were corrected in the second the effect may be regarded as negligible as the groups tested were small and control was complete at all stages.

iii.No group of children was tested twice on the same

day. The doing of the test should be an enjoyable experience if the claim ~~that~~ which Spearman makes that personality defects have little effect on test performance is to be upheld.The application of the test in one morning as was done in an Australian investigation which will be referred to later makes , in my opinion , the testing an arduous business for the subjects,especially the younger ones.A consideration of the order of difficulty of the sub-tests is interesting in this connection.

Table II. Order of difficulty of the Sub-tests.

Tests	1	2	3	4	5	6	7
Scottish Sec. Chidren.	6	4	7	5	3	1	2
Perth Central School Children	6	4	7	5	2	1	3
Perth Elem. School Children	6	3	7	5	2	1	4

See (13) Page 46.

The low ranks shown by the second group of sub-tests (Sub-tests 5,6and 7) show that it while shorter than the first group is ~~harder~~more difficult. This second group will therefore be more fatiguing than the first so it is desirable,if possible,that the subjects should be as fresh as possible. I therefore decided where possible to place half a school day (three periods) between the first and second groups of tests.If a class was given the first group in the second period of the day then the second group was given in the fifth period of the following day.and so on. This ensured in an approximate fashion that no form was favoured in the times at which it had to do the test.Early in the morning and towards the end of the school day are not regarded as good times for testing and while it is not possible to say that the results were not affected by the times of testing the children never lost interest and enjoyed doing the test very much indeed.

iv. The children were told that no marks would be given or taken away for spelling and neatness. This information in the primary standards particularly was always received with pleasure.

v. At the beginning of each sub-test the subjects were told not to begin writing until the complete question had been

given. This was done to prevent the lessintelligent from arriving at the answer by watching which alternative was chosen by the brighter children.It was found in the trial test that one or two subjects listened to the question and when they heard the word they thought was the correct one they wrote it down without waiting to hear the rest of the question.

vi. The timing of the pause between the questions was donecwith an ordinary pocket watch . Spearman suggests that the tester practise giving an eight second interval ~~without a watch~~ and then give the test without using a watch.

~~vii~~.No other modifications were introduced and great care was taken to see that the children understood the instructions before each sub-test was given.The procedure was made"as automatic as possible " as Spearman recommends.(17).Page 5.Before the testing was started I read the test out aloud several times so as to become phonetically accustomed to it.

2.Schools Visited .Arrangements were made with the headmasters and headmistresses s veral weeks before the testing commenced. These schools are classified in Table III.

Table III.Schools Visited.

Boys' Schools	Girls' Schools	Co-educational schools
1.Kingswood College	7.Diocesan School for Girls.	10. Grahamstown Primary School.
2.St. Andrew's Prep. School	8.Victoria Girls' High School.	11.St.Jpseph's Primary School.
3.St. Andrew's Coll.	9.Convent of the Assumption.	12. St. Peter's Primary School.
4.St. Aidan's Coll.		
5.Graeme College		
6. School of the Sacred Heart.		

All the testing was done in little over a month and the arrangements were excellent , everybody being most helpful and interested. Teachers in Grahamstown are accustomed to this sort of interference with their work . To what extent this sample of children is "test wise " can not be ascertained from the results of this investigation.

Teachers were asked to introduce me simply as " This

is Mr Butler. He is going to give you a test this morning in connection with work he is doing at Rhodes." At one of the girls' schools it was felt that the girls might behave foolishly and so a teacher remained in the classroom. When this happened I asked the teacher to stay in front of the class with me as I was afraid that the children might be embarassed if one of their own teachers saw what they were writing. There was no disciplinary trouble at all ; the children enjoyed doing the test and found it a very welcome break from their normal school work. Adolescent boys who often regard a new master , especially if he is still a student, as " fair game" and feel that enthusiasm for scholastic matters is beneath them showed great interest and gave little trouble. In only two standards was there any difficulty in getting the test started but once the work begun they settled down and enjoyed doing the test. In the secondary standards the girls were far more anxious about their results than were the boys but in the primary standards there was little, if any, difference between the attitudes of boys and girls.

The scripts were marked using a stencil and the scores right, wrong and omitted entered in a simple card index system. Other information as regards age, sex and general school performance was collected from the schools after the testing was completed.

A small number of Afrikaans - speaking children were also tested but their results were not included in the determination of the norms. Obviously a test given in a language that is only known inadequately by a child will not give a fair measure of the child's ability. The criterion used to judge whether a child was to be considered Afrikaans - speaking is the same as that used in the standardisation of the South African Group Test. If a child used Afrikaans better than he used English then he was regarded as Afrikaans - speaking. (See Page 65).

3. Results.

A. Determination of Norms.

i. Age Distribution. 1362 children were tested .An Age distribution was drawn up by placing those who , on the ~~first of~~ 1st April 1947 were 8years but not yet 9years old in the 8 year old class.and so on for the other ages. The class values are thus 8 years 6 months , 9 years 6 months , etc.These classes will be referred to as the 8.6 group , 9.6 group etc. Table ~~four~~ IV shows the distribution in the various age groups of both English and Afrikaans - speaking children.

Table IV. Age distribution of English- and Afrikaans-Speaking groups.

Age Groups.	Boys.		Girls.	
	Eng.	Afr.	Eng.	Afr.
7.6	1	0	0	0
8.6	8	0	2	0
9.6	41	0	37	1
10.6	89	5	53	4
11.6	114	4	70	3
12.6	109	4	77	6
13.6	145	9	77	6
14.6	172	8	85	4
15.6	94	8	56	3
16.6	38	2	15	4
17.6	5	1	3	1
18.6	1	1	0	0
Totals	817	42	475	28

Grand Total 1362.

It appears from this table that Grahamstown is primarily an educational centre for boys.as the number of boys is almost exactly double that of the girls in most of the age groups.As no forms higher than the Junior Certificate or its equivalent were tested the 15.6 and older groups are selected as superior children of these ages are in forms higher than the Junior certificate.(The terms "form" and

"standard" are used interchangeably.) The same applies to the age groups below the 9.6 group. but here it is the inferior cases that are omitted. The higher age groups are thus ~~superior~~ inferior and the lower superior from t his point of view.

The small number of Afrikaans speaking children is indeed surprising but the following facts must be taken into account :

a. Grahamstown is predominantly an English speaking centre from both rural and urban points of view.

b.All the schools with one exception are English medium schools.At one there are English medium classes* and only these were tested.These classes are formed using the same language criterion used here.The children at this schbol come largely from lower income group homes and it is probable that a large proportion of their number do not go on to the secondary standards . Consequently the number of children regarded as Afrikaans - speaking is lower than it would be if the children in the Akrikaans medium classes went on to the secondary schools. The primary departments of the other schools are use the English medium of instruction so that by the time children in these schools reach Standard III they have had four and often five years of instruction through the English medium.Those that aquire by this stage equal proficiency in English and Afrikaans would not be classed by their teachers as Afrikaans - speaking.

c.In line with the above statement as to the effects of instruction through the English medium we would expect the number of Afrikaans speaking children to decrease as we go up the age scale but this has not happened here.Many children come to school from outside areas where they have received their edu primary education through the Afrikaans medium. How many children we would have in the Afrikaans speaking group if there were parallel classes in the schools here is impossible to say but it would be much larger.It is,however,significant that at no school except the Grahamstown Primary School (Mentioned above) are there enough Afrikaans- speaking children to justify the formation of parallel classes.

The number of Afrikaans - speaking children , classified on the basis outlined above , represents 5.4 % of the total number.In

In 1930 Frances Bankes conducting a similar investigation here found the number of Afrikaans speaking children represented 6% of the total number.(3).

ii. Score Distribution.

A frequency distribution of the scoreswas drawn up and is given in Table V.This was done using a classifier (See Page 68). The class limits in the original sorting were 159.5 - 149.5 154.5 149.5 , 144.5 etc . The number of classes was then halved giving class limits as in the table below. with limits 159.5 , 149.5 , 139.5 etc.

Table V Frequency distribution of scores of Boys and Girls in the 9.6. to 14.6 age groups inclusive.

Class values	Boys	Girls	Total
154.5	7		7
144.5	81	29	110
134.5	154	60	214
124.5	123	70	193
114.5	110	70	180
104.5	61	66	127
94.5	67	41	108
84.5	28	27	55
74.5	28	24	52
64.5	7	9	16
54.5	3	3	6
44.5	1		1
	Grand Total		1069.

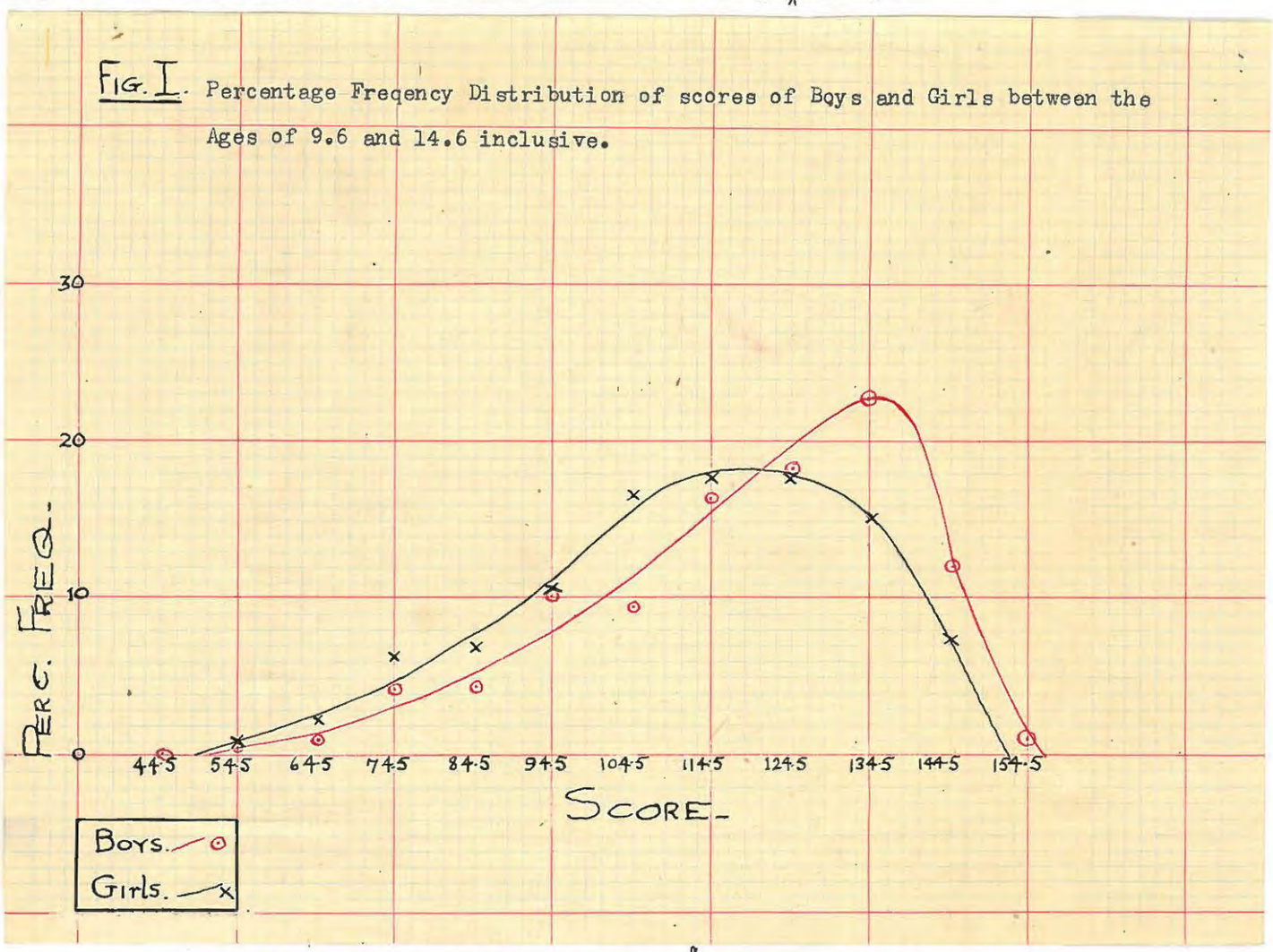
iii. Differences between the distributions.

Table VI gives the characteristics of the two distributions.

Table VI. Characteristics of the total score distributions.

	Mean	Median	Std. Dev.	Skewness	P.E.M.	P.E.Md.
Boys	117.8	121.0	20.88	-.45	.58 .42	.72.
Girls	111.7	113.7	19.48	-.28	.70	.84

Figure I shows the differences between the distributions graphically.



To establish whether there ~~is~~ are significant differences between the distributions the probable errors of the differences between the means and the differences between the medians were calculated from the probable errors given in Table VI.

i. The differences between the means was found to be 6.1 with a probable error of .91. This difference is nearly seven times the probable error and is thus clearly significant. Holzinger says that a difference that is more than four times its probable error may be regarded as significant.(11). Page 237.

ii. Medians. Here the difference is 9.3 and the probable error 1.107, again a clearly significant difference.

iii. Standard Deviation. This difference is 1.4 and the probable error .61 . This difference is not therefore significant. so that as regards dispersion these distributions are not significantly different.

iv. Skewness. The negative skewness is more marked in

the case of the boys.A reason for this negative skewness can be found in the work of Hardie and Bankes.(3)/and(10-)/Both find insufficient Page 59 Page 191 scatter of scores at the top of the age scale.Spearman says that many fifteen year olds will find the test too easy and as many of the cases here are almost fifteen years of age we would expect a negatively skewed curve. But this does not explain the greater degree of skewness on the part of the boys.If we assume that there is no significant difference between the performance of boys and girls on this test (Marshall using the results from testing 6843 children found no significant difference) then this weakness of the test in showing insufficient scatter of score in thevupper age groups should have the same effect on boys and girls.Although differences between the performances of boys and girls have been found to exist most standardisations are made on the assumption that there is no significant difference. If the above assumption is ~~accep~~ accepted it would appear that an appreciable sampling error is responsible for the significant superiority of the boys.~~here~~ It seems likely that there is a good group within the total group of boys that is causing this concentration of scores at the top of the age scale.

Having established that the two groups are significantly different the next step that appeared fruitful was to see whether there was any significant difference between the median scores of the corresponding age groups.

Table VII . Median scores for Boys and Girls in specified age groups.

Age groups	Boys		Girls	
	Medians	P.E.	Medians	P.E.
9.6	92.68	2.66	94.50	2.30
10.6	97.10	1.66	101.4	2.04
11.6	115.50	1.395	102.58	1.99
12.6	122.90	1.48	116.01	1.805
13.6	128.60	1.08	125.30	1.58
14.6	132.70	.86	126.85	1.44.

The boys scores are higher from the 11.6 group

onwards.The reliabilty of the differences was tested as before.

Table VIII Reliability of the differences between Boys' medians and Girls' medians.

Age groups	Difference	P.E. of Difference.
9.6	1.82	3.52
10.6	4.3	2.63
11.6	12.92	2.43
12.6	6.9	2.33
13.6	3.3	1.91
14.6	5.95	1.67

Only the difference between the 11.6 medians is significant. Here,however,the probable errors of the medians are all larger as the number of cases from which the medians are calculated is much smaller.Hence to establish the significance of the difference between medians (means) in each age group a larger difference would be required in each case.This,while it might be theoretically possible,is extremely unlikely. Even if the differences were larger they would still have to be sufficiently large to be significant. In any case in each age group the number of cases is small and the probable error will have little significance.The superiority of the boys' group may be regarded as proved from the consideration of the ~~tests~~ total distributions.

We have established that the boys' group is as a whole superior.We may suggest an explanation by reference to the list of schools tested.(Table III.) The first four on the list Kingswood College, St. Andrew's Preparatory School , St. Andrew's College and St. Aidan's College are private schools at which the fees are very high indeed.being on the average about £150 a year. The large majority of the boys in these schools must come from upper income group homes. We would expect them to be a superior group ; the continuous gradation of intelligence as measured by tests through the various income groups is well established.

Let us therefore divide the boys into two groups, Good and Poor,

on this basis in order to test whether there is any significant difference between our arbritarily defined groups. We employ the same line of reasoning used to show the superiority of the boys over the girls.

iv. Differences Between Good and Poor groups .Boys.

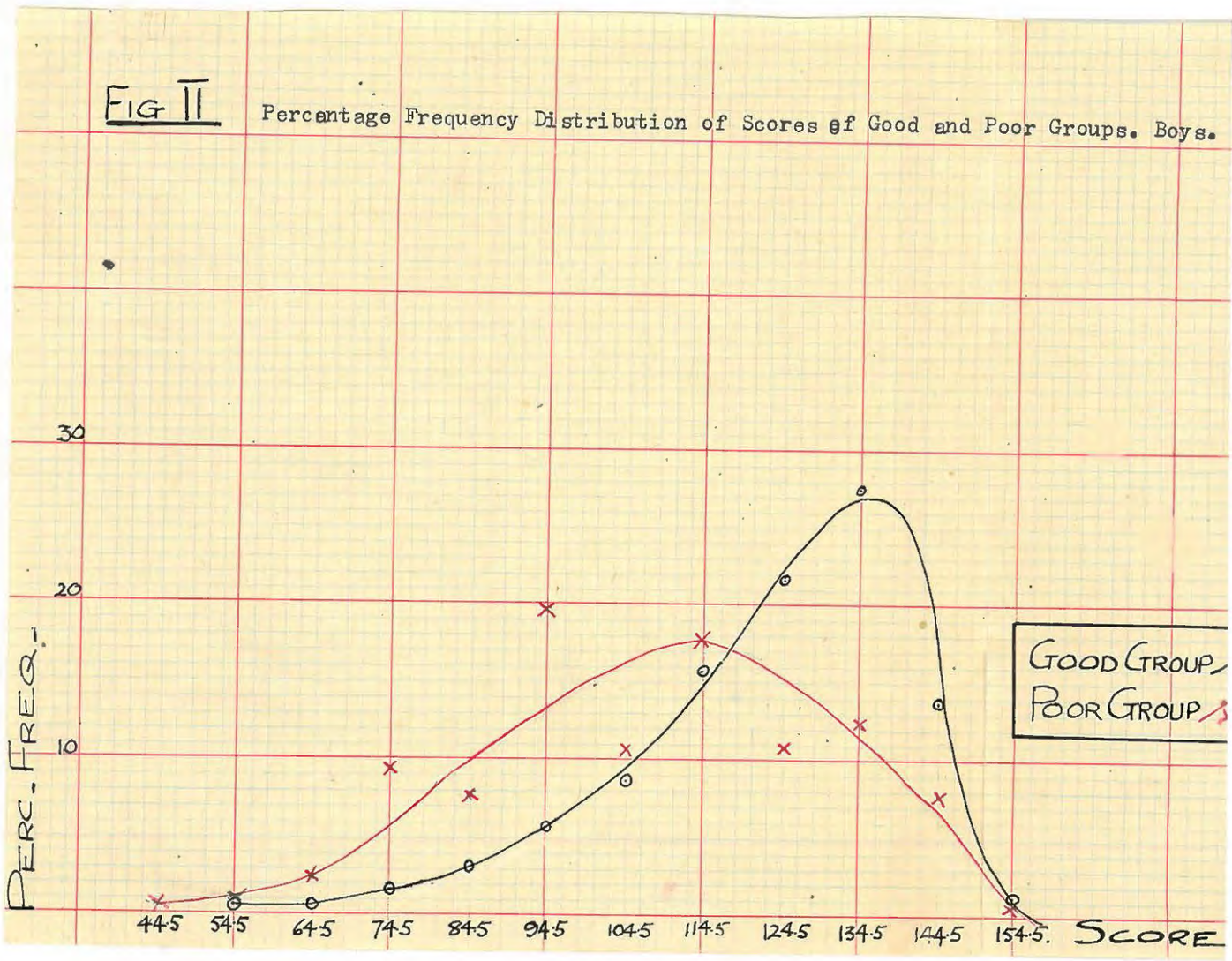
Table IX. Numbers in specified age groups of Good and Poof groups.Boys.

Age groups	Good group	Poor group	% Good group.
9.6	25	16	61
10.6	53	36	60
11.6	79	35	69
12.6	77	32	71
13.6	103	42	71
14.6	130	42	76
Totals	467	203	

Table X Score distributions of Good and Poor groups.Boys.

Class values	Good group	Poor group
154.5	6	1
144.5	65	16
134.5	129	25
124.5	101	22
114.5	74	36
104.5	40	21
94.5	27	40
84.5	12	16
74.5	9	19
64.5	2	5
54.5	2	1
44.5		1
Totals	467	203

A percentage frequency distribution was drawn up and the curves drawn as shown in ~~Tab~~ Figure II.



A free hand curve drawn from a distribution as irregular as that shown by the poor group must be interpreted with extreme caution. The highly skewed curve shown by the distribution of the scores of the good group is clearly shown by the graph. No matter how the curve representing the scores of the poor group is drawn it will not approximate remotely that of the good group.

The characteristics of these distributions were calculated.

Table XI. Charateristics of the score distributions of Good and Poor groups. Boys.

	Means	Medians	Std. Dev.	Skewness	P.E.M.	P.E.Md.
Good group	122.85	126.18	16.63	-.60	.52	.65
Poor group	107.65	108.78	22.48	-.15	1.06	1.33

The significance of these differences was tested as before .

i.Means. Here a difference between the means of 15.2 in favour of the Good group was found.The probable error of the difference was 1.18 making the difference clearly significant.

ii.Medians.A difference of 17.4 also in favour of the Good group with a probable error of 1.48 was found ; again clearly significant.

iii. Standard deviation. The difference is 5.85 and the probable error .83 ; also significant.

iv. Skewness.The Poor group scores are closer to zero skewness than the Good group scores.The high negative skewness and the smaller standard deviation of the Good group suggest that the scores of the Good group are concentrated at the top of the score scale.This is clearly brought out by the percentage frequency distribution curve.

Our hypothesis is borne out by the results as the differences are clearly significant.Age group medians were again calculated and the differences between the respective medians calculated. Probable errors were not calculated as the number of cases is too small. We are not basing our argument on the differences between the medians in each age grpup but on the differences between the characteristics of the total distributions of Good and Poor group scores..The medians do , however , show large differences in favour of the good group.

Table XII. Medians of specified age groups of Good and Poor groups .Boys.

Age groups	Good group	Poor gfoup	Difference.
9.6	93.4	89.5	3.9
10.6	105.75	87.8	17.95
11.6	118.5	107.0	11.5
12.6	126.1	102.0	24.1
13.6	131.4	115.2	16.2
14.6	133.3	129.5	3.8

The medians of the poor group although irregular are all below those of the Good group .

The usual procedure in standardisations of this type where an unlimited number of cases is available is to so choose the schools that upper and lower income groups are equally represented. This was not possible here as the small number of cases available made is necessary to use all the material . As we intend to derive norms that will be applies to girls as well as boys we are working on the assumption that boys are not on the average superior to girls. Our object is, therefore , by purposive selection to eliminate the superiority of the boys which may be ascribed as due to the nature of our particular sample.

V. Differences between Good and Poor groups .Girls.

Only one girls' school , The Diocesan School for Girls , can be classed as "Good" on the same basis as used in the discussion of the boys' schools. It does , however , show a consistent superiority over the other girls' schbols.

Table XIII. Age distribution of Good and Poor groups.

Girls.

Age groups	Good group	Poor group
9.6	8	29
10.6	13	40
11.6	14	56
12.6	18	59
13.6	22	55
14.6	16	69
Totals	91	308

In all the age groups the number of the Poor group is more than three times the number in the Good group.

A score distribution of Good and Poor groups was drawn up.

Table XIV. Score distributions of Good and Poor groups .Girls.

Class values	Good group	Poor group
154.5		
144.5	13	16
134.5	21	39
124.5	18	52
114.5	14	56
104.5	11	55
94.5	9	32
84.5	4	23 ~~32~~
74.5		24
64.5	1	8
54.5		3
Totals	91	308

The characteristics of these distribution were calculated as before.

Table XV . Characteristics of total score distributions of Good and Poor groups .Girls.

	Mean	Median	Std. Dev.	Skewness	P.E.M	P.E.Md.
Good group	120.44	123.1	18.09	-.44	1.28	1.60
Poor group	109.17	111.11	20.90	-.28	.80	1.007

No time need be wasted with a detailed consideration of these figures . The good group is clearly superior. It is interesting to note that the same tendencies are shown by both the Good groups. Means and medians are significantly greater , they show a greater degree of skewness and the standard deviations are smaller.Our basis for the division of the schools into two groups seems to be justified from the consideration of both boys' and girls' score distributions.

Table XVI gives the medians of the age groups in Good and Poor groups . No probable errors were calculated here as the number

of cases is small.The differences are all in favour of the Good group and most are very large and might even be shown to be significant but with such a small number of cases it is extremely dangerous to arrive at conclusions based on the probable errors.

Table XVI. Medians of specified age groups of Good and Poor groups .Girls.

Age groups	Good group	Poor group	Difference
9.6	102.0	88.6	11.4
10.6	104.5	100.5	4.5
11.6	112.8	101.8	11.64
12.6	129.5	113.56	15.94
13.6	133.5	121.3	12.2
14.6	139.5	125.75	13.75

The reason for the superiority of the boys is the fact that more Good boys are included than Good girls . It would seem likely that the differences would disappear if the Good girls were ~~rep~~ represented in the same proportion as the Good boys.This group of girls make up only 25% of the total number of girls while the corresponding boys' group constitutes 68% of the total group of boys.

vi.Correction of the distribution. It was decided to eliminate half the number of cases in the Good group of boys from the 11.6 group upwards.This elimination would have the effect of giving roughly equal representation to Good and Poor groups.and was done by placing the cards in alphabetical order in age groups and eliminating every alternate card.This would give a group that is representative of the Good group but reduces the number of cases by half.The rationale of this decision is as follows :-

a.The sample of boys is a superior one as shown by Tables VI ? VII and VIII and Figure I.

b.This superiority of the boys is due to the large proportion.of the Good group as shown in Tables IX , XI and XII.

c. This elimination would have the effect of :-

1.Giving roughly equal representation to Good and Poor groups.

ii.Bringing the number of ~~girls~~ boys to approximate equality with the girls.

iii.Reducing the negative skewness of the total distribution of the scores of the boys relative to that of the girls.

d.The sample of girls was left untouched because while the Good and Poor groups could have been given equal representation by the elimination of two thirds of the Poor group this would have resulted in too small asample.The Poor group of girls shows a slight superiority over the Poor group of boys which suggests that the distinction between Good and Poor groups does not hold with equal force in this group.Grahamstown schools are mostly schools for boys and it is quite prbbable that a number of school girls at present in the Poor group would be in the Goodgroup if the facilities were available.

The sample of boys from a purely statistical standpoint is now the one most likely to give representative norms.The sample of girls must be regarded as inferior from this point of view but this approximation is the best that can be achieved with the material available.

The number of cases eliminated was 196 out of a total of 670 leaving a remainder of 474,193 cases of the Good group out of an original 389 in the age groups concerned in the elimination remain ; the elimination began with the 11.6 group as stated before.This point was chosen because from this group onwards the numbers in the Good group are more than twice those in the Poor group.

Table XVII. Numbers in Good and Poor groups after the elimination of half the number of cases in the Good group from the 11.6 group onwards.

Age groups	Good group	Poor group	Good	Total	Boys	Girls
9.6	25	16	61	41		37
10.6	53	36	60	89		53
11.6	39	35	53	74		70
12.6	38	32	54	70		77
13.6	51	42	55	93		77
14.6	65	42	61	107		85

The elimination has had the desired effect. It might well be argued that the elimination should have become progressively h heavier as we went pp the age scale. This could well have been done but the effect would have been small. It will noticed that the ~~di-ffe~~ difference between the medians of the 14.6 groups as shown in Table XII ~~is) (i) (s) p)~~ is the smallest of all due presumably to the falling - off ~~of~~ the efficiency of the test in this and higher age groups. The boys are over represented in the 13.6 and 14.6 groups . Equality in numbers is not essential if there is no significant difference between the two groups. (~~See page~~).

The distributions now have the following characteristics.

Table XVII. Characteristics of score distributions of boys and girls after the correction of distribution of boys' scores.

	Mean	Median	Std. Dev.	Skewness	P.E.M	P.E.Md
Boys	114.75	117.52	19.55	-.42	.61	.76
Girls	111.7	113.7	19.48	-.31	.66	.82

The elimination has had the effect of bringing the skewness of the boys distribution closer to equality with the ~~boys~~ girls. A percentage frequency distribution was calculated and the results plotted giving the curves as shown in Figure III.

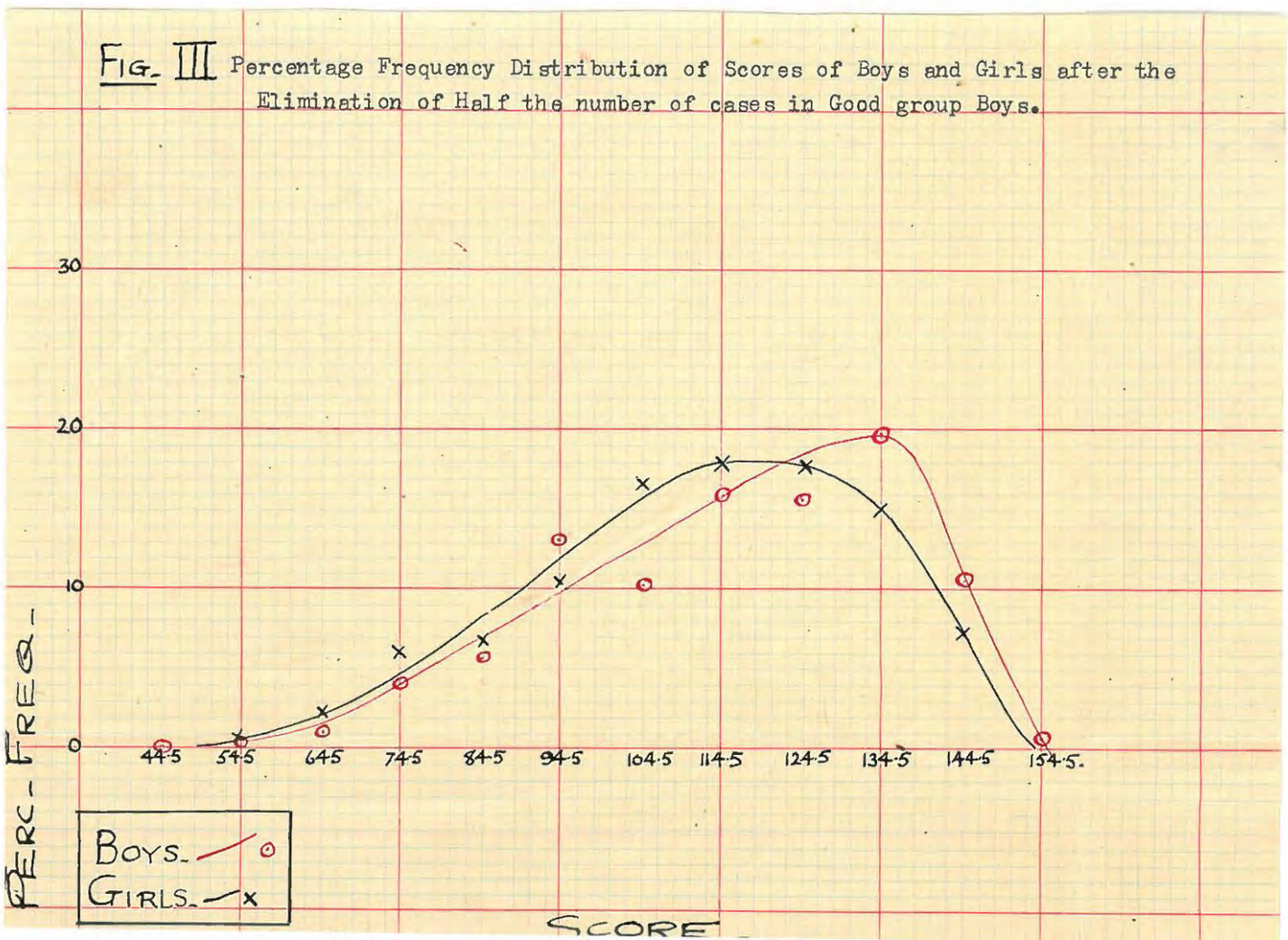


FIG. III Percentage Frequency Distribution of Scores of Boys and Girls after the Elimination of Half the number of cases in Good group Boys.

The differences between the boys' and girls' groups as they now stand were next considered.

Table XIX. Reliability of the differences between the means and medians of boys'group and girls'group.

	Difference	P.E. Diff.
Means	3.05	.90
Medians	3.82	1.12

Neither of these differences is significant.

The medians of the age groups of boys'and girls' distributions are shown in Table XX with their probable errors.

Table XX. Medians of boys'and girls' distributions in age groups

Age group	Median. Boys .	P.E	Median.Girls	P.E.
9.6	92.68	2.66	94.5	2.30
10.6	97.1	1.66	101.4	2.04
11.6	113.07	1.898	102.58	1.99
12.6	120.08	1.99	116.0	1.81
13.6	126.06	1.45	125.3	1.58
14.6	133.1	1.20	126.8	1.44

The differences between these medians and their probable errors are given in Table XXI.

Table XXI Reliability of the differences between the medians of boys' and girls' distributions in age groups.

Age groups	Difference	P.E. Diff.
9.6	1.95	3.52
10.6	4.3	2.63
11.6	10.49	2.75
12.6	4.08	2.69
13.6	.76	2.15
14.6	6.3	1.87

By the same reasoning employed in the comparison of the differences between the uncorrected distributions we arrive at the

conclusion that the characteristics of the corrected distributions are not significantly different and we can therefore combine these two distributions for the purpose of calculating norms for each age group. In the above table it will be noticed that only one of the differences that of the 11.6 group is significant.

The distributions were combined and medians calculated.

vii. Calculation of the combined norms.

Table XXII. Medians derived from combined distributions.

Age groups	Medians	P.E.
9.6	93.3	1.72
10.6	98.8	1.23
11.6	107.9	1.44
12.6	117.8	1.28
13.6	125.6	1.07
14.6	131.1	.93

To establish whether that the differences between the medians of the different age groups are not due to sampling errors we test the reliability of the differences between the medians as in Table XXIII below.

Table XXIII Reliability of the differences between the medians.

Age groups	Differences	P.E. Diff.
9.6 - 10.6	5.5	2.12
10.6 - 11.6	9.1	1.90
11.6 - 12.6	9.9	1.93
12.6 - 13.6	7.8	1.67
13.6 - 14.6	5.5	1.42

6

The difference between the 9.6 and 10.6 medians is probably not significant and this is due to the fact that the 9.6 group must be considered a selected one as the inferior nine year-olds are in forms below those we have tested. The 9.6 norm is therefore higher than it would be if all the nine-year-old children in Grahamstown

had been tested.This norm cannot be considered as reliable.In the same way the effects of selection may be held responsible for the insufficient difference between the 13.6 and 14.6 norms.Superior children in the fourteen year - old group are in forms higher than those tested. and so the 14.6 norm is lower than it should be. On the other hand the differences between the others are significant which is inconsistent with the hypothesis that differences are due to sampling errors.This suggests that the difference in score are due to differences in age.

Using the method of averages a curve was fitted to these medians .The 11.6 median was taken as the origin and only the 10.6 ,11.6 , 12.6 and 13.6 medians used in the determination of the equation.The calculated values are almost identical with the observed ones while the values obtained for the 14.6 and 9.6 groups are closer to what one would expect from a representative sample.No reliance can be placed on these two values.

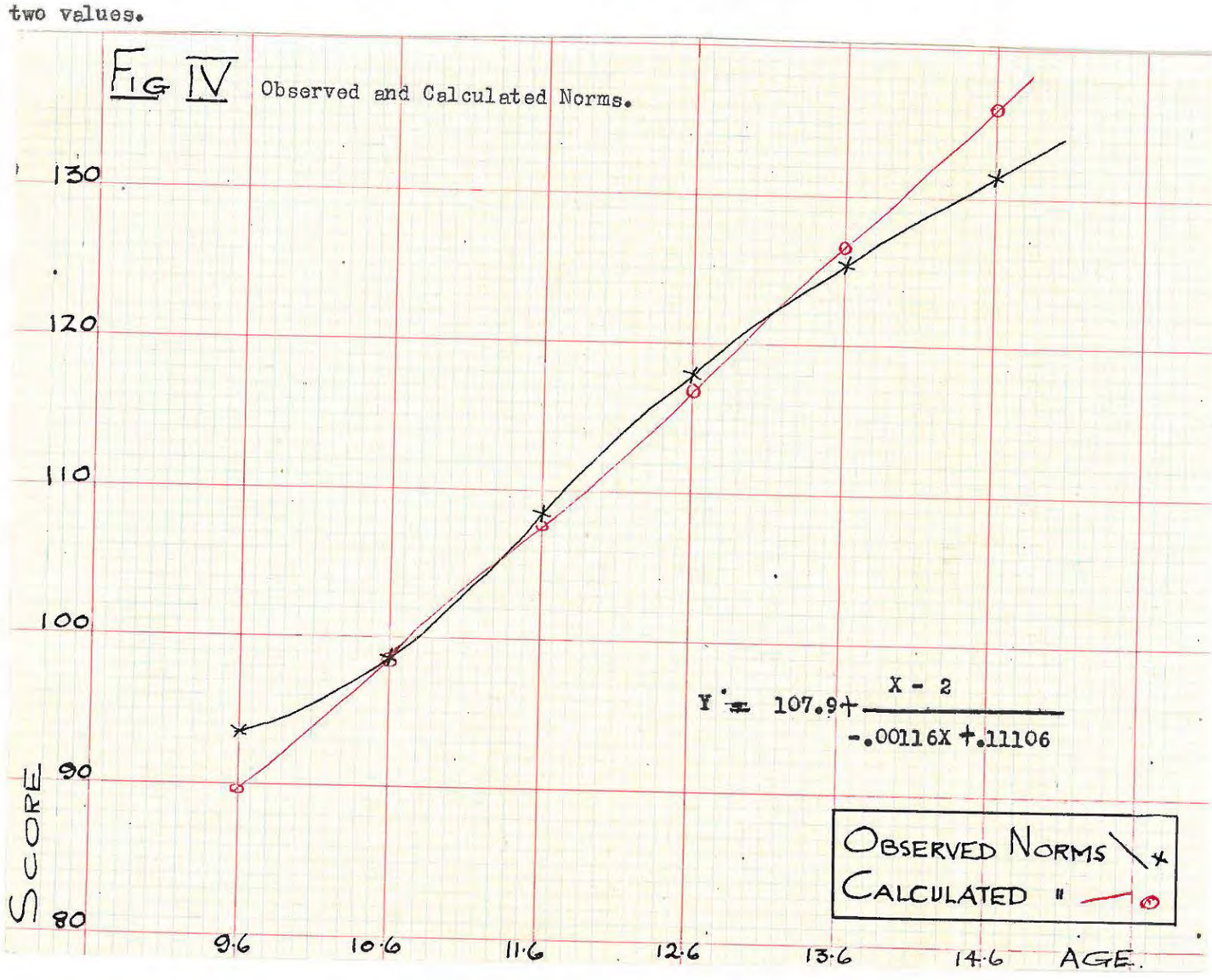


Table XXIV gives the observed and calculated values represented by these curves.

Table XXIV. Observed and calculated norms.

Age groups	Observed values	Calculated values
9.6	93.3	89.9
10.6	98.8	98.8
11.6	107.9	107.9
12.6	117.8	117.1
13.6	125.6	126.6
14.6	131.1	136.4

To summarise the method used in calculating these norms:

i. By comparing the total distributions of scores of boys and girls a significant difference was found between them.

ii. The medians for each age group showed a consistent but statistically insignificant difference in favour of the boys from the 11.6 group onwards. The difference between the 11.6 medians was significant.

iii. The boys were next considered in two groups. The Good group is a group of private denominational schools at which the fees are very high indeed. This fact leads us to assume that the children in these schools come from upper income group homes. This group was shown to be significantly superior to the Poor group.

iv. The Good group were found to be over-represented from the 11.6 group onwards.

v. To overcome the unrepresentative character of the sample half the cases in the Good group were eliminated from the 11.6 group onwards.

vi. After this elimination the boys showed a superiority over the girls that was found to be insignificant. Medians of age groups showed insignificant differences. This is not strictly apposite as the argument is based on the differences between the total distributions.

vii. The scores of the boys and girls were combindd and a median calculated for each age group. The differences between the

medians of different age groups were found to be significant except in the case of the 9.6 and 14.6 groups which are both selected to a certain extent.

viii.These norms were fitted to a curve by the method of averages .The differences between the observed and calculated values are small except in the cases of the 9.6 and 14.6 groups where the calculated values are closer to what one would expect them to be.No reliance can be placed on these values.

B.DETERMINATION OF THE VALUE OF THE INTELLIGENT QUOTIENT REPRESENTED BY THE CURVE OF NORMS.

As the schools did not have available information as to the incomes of the parents it was not possible to draw up a frequency distribution showing how the various income groups are represented in this sample.Such a distributionwould be of great value in deciding to what extent theis sample is superior to a truly representative ~~one~~ South African sample both in the determination of the norms and in the allied question of deciding what level of intelligence is represented by these norms.A frequency distribution of the incomes ~~is~~ earned in the Albany district would be of little value as so many children come here from other areas. The problem was therefore approached from another direction.

1.Test with the South African Group Intelligence.

Test.

An attempt to establish what intelligence quotient is represented by these norms was made by testing 82 children at Kingswood College,who had done the Spearman test , with the South African Group Test. From these results their intelligence quotients were calculated giwing the following results.

Median Age 12.6.

Median I.Q. 117.6 .

Median Score (Spearman test) 121.8.

This group of boys is a superior one as shown by the median I.Q.This group of boys should also be superior as regards this group as they come from one of the Good schools. If we assume that our curve represents an I.Q. of ~~100~~ 100 then ~~the median~~ a Spearman score of 121.8 represents an I.Q.of 104. This group is therefore only slightly superior as regards the whole group from which thesenorms are calculated but the whole group would appear to be superior compared to a truly representative South African sample.

The mental age of a child scoring 121.8 on the Spearman test may be found from the relation:

$$\text{Mental Age} = \frac{\text{Chron .Age} \times \text{I.Q.}}{100}$$

$$= \frac{12.5 \times 117.6}{100}$$

= 14.7 years.

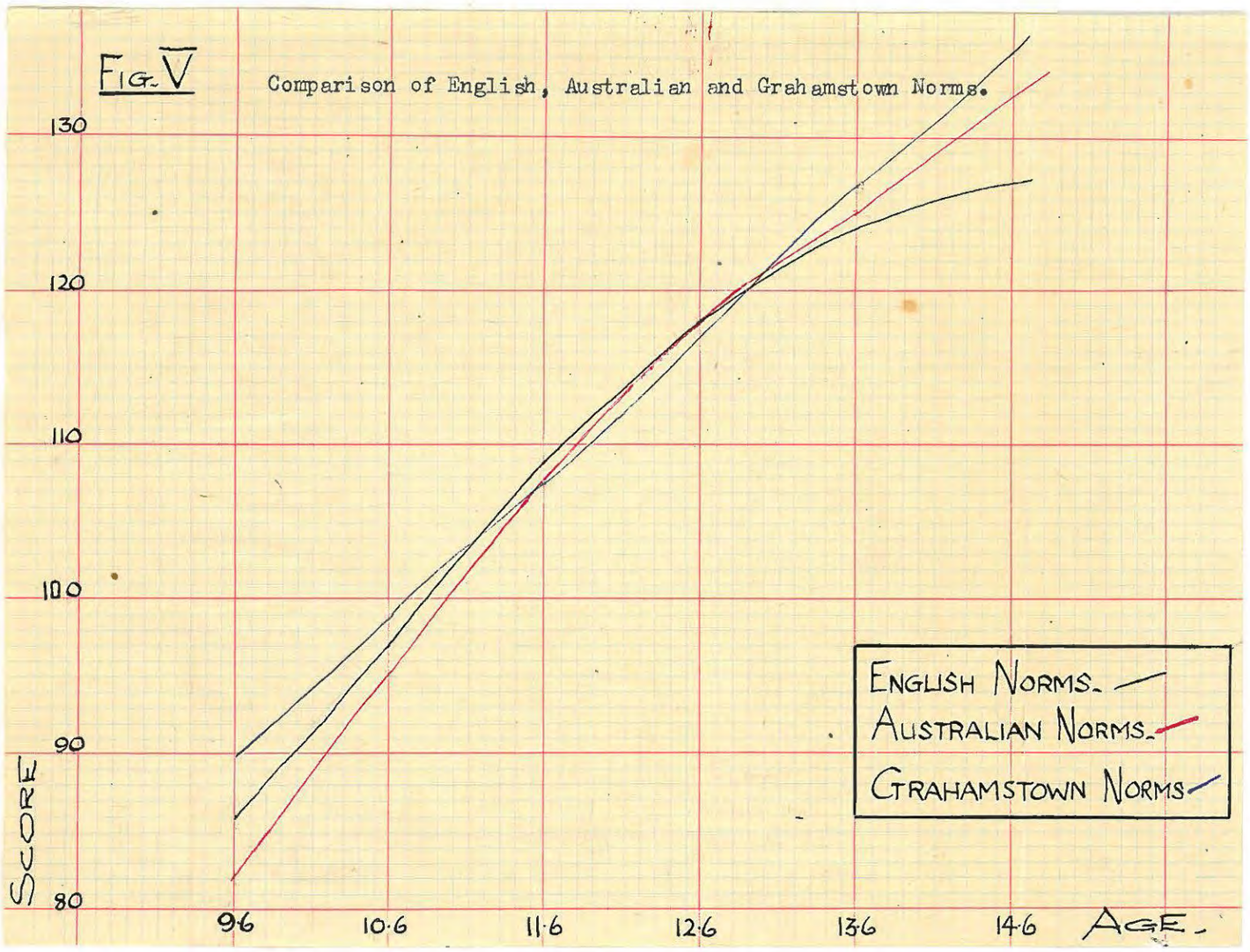
As stated above this group of 82 would have an I.Q. of 104 on the basis of the Spearman test results if we assumed that our norms represent an I.Q. of 100. But the South African Group Test shows to them median to have an I.Q. of 117.6. Therfore the I.Q.'s shown by these curves are too low and the correction factor is approximately $\frac{117}{104}$ which equals 1.13. i.e. in order to get an approximate value of the I.Q. as shown by the South African Group Test we would have to multiply our I.Q.'s by 1.13. This result is not accurate enough in my opinion to warrant the drawing up of a scale of I.Q.'s to represent the South African population as a whole. These norms must be taken to apply to this area only ; any further generalisation would be extremely dangerous. A better way of establishing the position of the curve relative to that which would be shown by a representative sample would be to have a sample of , say one hundred, children tested with an individual test and their I.Q.'s accurately found. Two conclusions emerge :

1. The I.Q.'s given by these norms would have to be multiplied by 1.13 to approximate to those that would be given by the South African Group Test when used on the same children.

2. The sample of children on which these norms are based is superior when compared to the South African population as a whole.

2. <u>Comparison with English and Australian norms.</u>

In both the English and Australian standardisations each question was given twice whereas as here the questions were given once only except in the first two tests where they were given twice. This method of presentation is one that we would expect to make greater demands on the subjects. The effect of this method can be judged from a consideration of the curves representing the Australian , English and the norms calculated for this area.



There is remarkable degree of correspondence between these three curves. If we assume that South African children in general are as intelligent as Australian children then this sample is a superior one as it has given norms which are similar to the Australian even though the test was given in an manner which we may assume is less advantageous. Continuing with this assumption we are able to judge what the effect of this method of presentation has been. Using the factor found above a score of 117.8 given by a child of 12 years 6months represents ~~a score~~ an I.Q. of 114.5. On the Australian norms a child with an I.Q. of 114.5 would obtain a score of 127, a difference of 10 points.

This method has had the effect of introducing a reproductive element into the testing in the sense that subjects have to remember the question sufficiently well to be able arrive at the solution after only one presentation. Whether this has decreased the diagnostic val ue of this test is difficult to say nor is it possible to say whether

all ages are affected equally.All the subjects are affected by it here so individual differences will still be shown.The high raw score correlation between this test and the South African Group Test seems to show that it has been little affected.In this test the reproductive element is removed by the provision of printed test forms. Whatever the effect of this method of presentation it has not affected the variation of performance with age.The correspondence with the two overseas curves is remarkable and would seem to suggest that the diagnostic value has not been impairdd. Spearman would not have found this result surprising as he claims that there is no factor which can be isolated that applies to the method of presentation .Even though this test is a superior sample the curve shows that variation of performance with age is practically identical with that of the other two standardisations referred to.

(See page 58)

C. VALIDITY. 1.Correlation with general school performance.

The test was correlated in form groups with general school work measured by the term averages of the children.The number of cases in each group is small (never more than forty) so the correlation coefficients taken singly have little quantitative significance.The cons consistently high and positive values are however an indication of validity.

Table XXV.Distribution of correlation coefficients .

Class limits.	Freqency
.995 - .895	1
.895 - .795	3
.795 - .695	5
.695 - .595	9
.595 - .495	12
.495 - .395	9
.395 - .295	5
.295 - .195	6
.195 - .095	2
.095 - 0.	1
Total	53

There were three negative coefficients all from forms of boys ; the lowest value was between -.095 and -.195. The median value of this distribution is+.52.No probable error is given as it would have no meaning.

2. Correlation with South African Group Test.

The raw scores given by the 82 children testdd with the above test and the Spearman test were correlated. This gave a result of +.73 $\pm$ .034.This result has considerably greater significance than thatof the preceeding section but the number of cases is still not large enough . Correlations can be misleading when the number of cases is less than 100.

3.Other results.

i.Marshall obtained a correlation of +.78 $\pm$ .037 between the Binet test and the Spearman test.and one of

+.85 ±.008 with the Sidney Teacher's College Group Scale.The number of cases was 50 and 612 respectively. (13) Page 36.

ii.Frances Bankes correlated the Spearman test with the Grey scale and the South African Group Test and obtained correlations of+.77 ±.0165 and+.81 ±.0185 respectively.She used only half the number of questions in the Spearman test, (3).Page 29. No. of cases 350.

iii.Hardie correlating general school work with results on this test obtained a correlation of +.412 ±.018.which when corrected for attenuation gave a correlation of+.724. Hardie's method of estimating the standard of general school work was superior to the one used here. A special examination was set and marked by him and as he did not knpw the children th- it is likely that his measure of their work was a more impartial one than that used here where the likes and dislikes of the teachers are bound to have had an effect. (10). Page 190.No. of cases 761.

D. SEX DIFFERENCES.

1.Norms.The consistently lower values shown by the girls in all age groups except the 9.6 and 10.6 groups may be regarded as due to a sampling error.Marshall's standardisation based on the results of testing 6843 Australian children found no significant difference between the two groups .~~It is~~ gThese high norms appear as we saw earlier to be due to the large number of boys from upper income group homes who come to school in Grahamstown.

2.Attitude. The girls as would be expected gave little trouble , were keener than the boys and showed greater anxiety about their performances.~~then di~~Boys often tried to waste time but once the test ~~begun~~ was started they became interested and their interest was maintained throughout. Adolescent girls were often embarassed and tended to giggle at first but this soon passed off and they settled down without further difficulty.

3. Four boys gave scores of over 154 but no girls did ; the only score below 49 was gained by a boy. This result is in agreement with the hypothesis that states that boys show a greater variability than do girls.

4. Correlation with school work. The correlatinns studied in connection with the validityof the test (See page 58) were treated to see if there was any difference between the two groups. The girls showed a median correlation of+.565 and the boys one of+.473. While there is no quantitative significance in this result it is what we would expect as girls are generally considered better students than boys and so consequently make more efficient use of their intelligence.

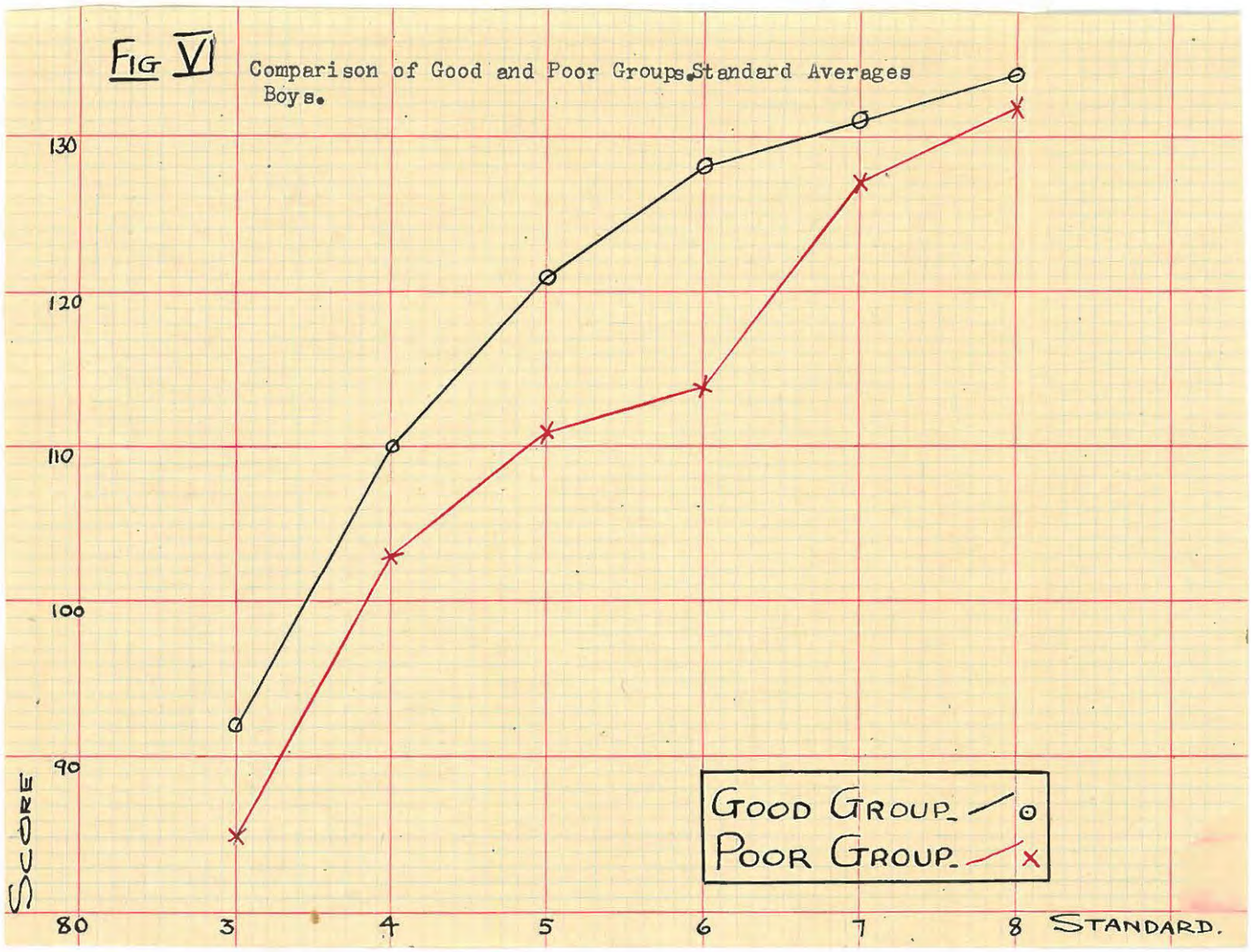
E. COMPARISON OF EQUIVALENT FORMS. Average scores for each form group were calculated. This treatment includes all the cases tested iee those whose results we have already considered in the determination of the norms and those who were eliminated as being either over - or under -age.

Table XXVI Form averages .Boys.

Form	Kin.Col.	St.A.Pr.	St.A.Co.	St. Aid.	Grae.	Sch of S.	Gr.Pr.	St.Pet.
3	100.6	83.6		96.3	91.47	79.6	77.7	77.8
4	114.2	109.2		106.5	106.5	102.5	92.0	100.8
5	121.2	121.6		121.9	111.8	98.8	123.0	98.6
6	125.4	138.2	127.9	120.8	115.3	116.5	104.5	117.0
7	132.1		131.9	131.1	127.0			
8	134.1		134.5	135.1	132.0			

(The above ~~list of school~~ abbreviations of the names of the schools may be ~~identified~~ interpreted by reference to Table III, page 33. They are numbers 1, 2 , 3 , 4 , 5 , 6 , 10 and 12 respectively.)

The first four schools in the above table are those we ~~el~~ included in the Good group. Their scores are all superior to those of the Poor group except in the case of the Form 3 score at St Andrew's Preparatory School and the Form 4 score at St Aidan's. The standard five score at the Grahamstown Primary School is the highest in that group but it must be remembered that this group is very small compared to to those in the other schools. If the results for Good and Poor groups are combined in each group and the results plotted on a graph the following figure results.



If we assume that the Form VI (Poor) group is subject to a sampling error then the curves are roughly "parallel". The Good group shows a consistent superiority over the Poor group as we saw before. There is a tendency for the differences between the groups to decrease as we go up the age scale due presumably to the falling off in the efficiency of the test at these higher ages. To consider this superiority from another point of view the average ages of the children in each formed in the two groups were calculated and compared. See Table XXVII overleaf.

Table XXVII. Average ages of forms in Good and Poor groups.

Form	Good group	Poor group
3	10.6	10.0
4	11.2	11.8
5	12.1	12.4
6	13.4	13.9
7	14.2	14.5
8	15.2	15.3

(The ages given above are expressed in years and months.)

In every case except that of Form 3 the Good group ages are below those of the Poor group. The higher scores of the Good group in the form averages are them more significant than at first appears.in

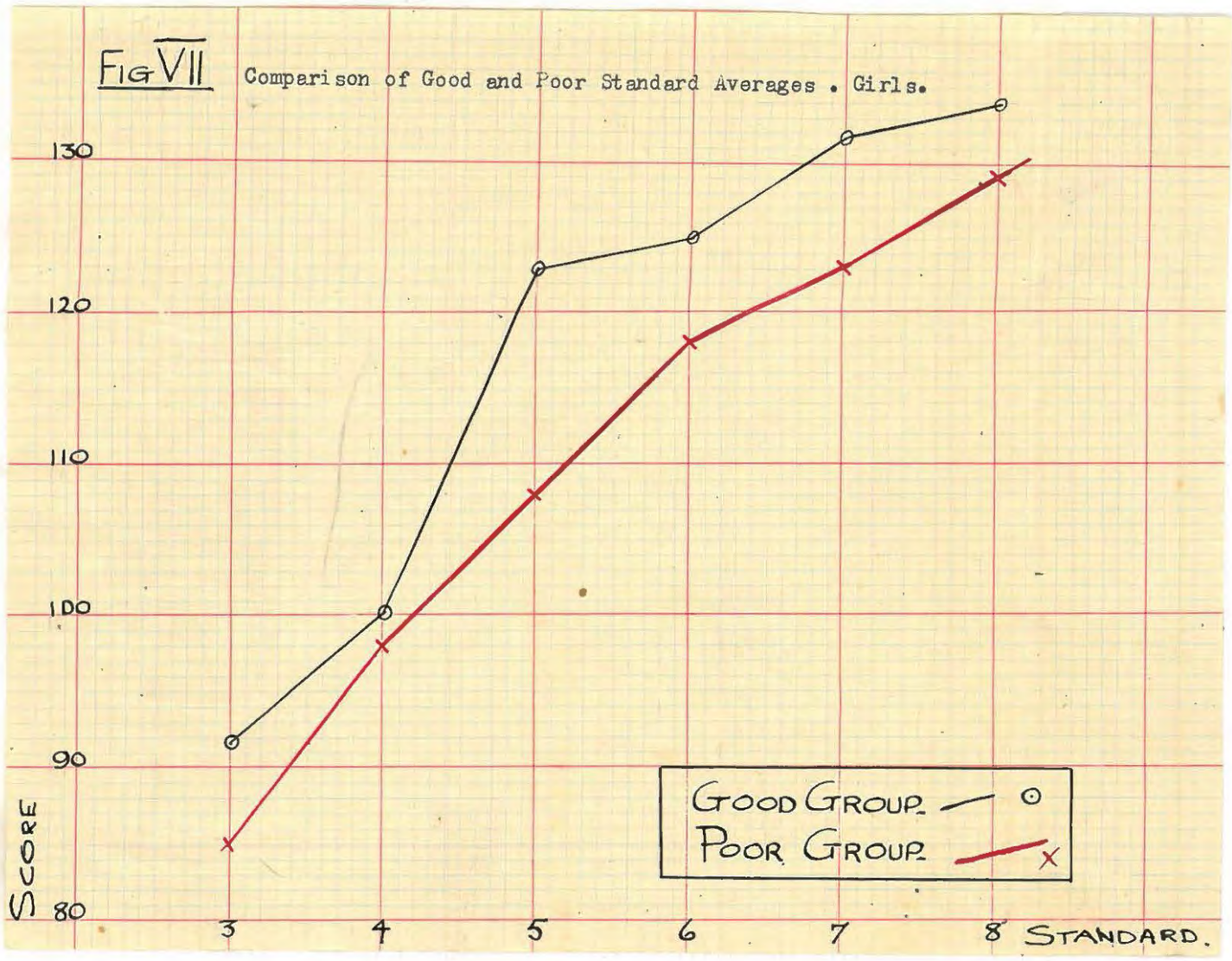
The form averages were calculated for the girls' schools giving the following results.

Table XXVIII Form averages .Girls.

Form	D.S.G	V.G.H.S.	Con.	G.P.	S.P .	S.J
3	101.5	87.1	93.6	97.6	79.2	82.4
4	99.6	100.8	103.0	86.0	92.5	93.2
5	122.8	112.0	110.0	102.2	98.6	110.5
6	125.0	116.7	122.5	111.3	109.0	126.0
7	131.8	122.5	126.7			
8	134.3	129.3	126.7			

(These abbreviations represent schools numbered 7 ,8 , 9 , 10 , 11 and 12 respectively.Table III Page 33.)

The only school that can be classed as Good on the same basis as used when considering the boys' schools is the Diocesan School for Girls , the first in the above table.Here also the average scores are with two exceptions above the rest. When the scores are combined in the Poor group the following curves can be drawn.



The same consistent superiority is shown here. The superiority of theGood groups appears therefore to be well established both from the point of view of the age groups and that of the school form groups.This,while it is not a new result , shows that the test is fufilling its functions in that it shows differences of ability which have been shown before to exist on the above basis.

F . AFRIKAANS CHILDREN .

Table XXIX. Medianscores and ages of the Afrikaans-speaking group.

	Median score	Median age
Boys	111.5	13.1
Girls	92.8	13.4

The effect of doing the test in aforeign language can be guaged roughly if we assume that the sample of Afrikaans boys is at the same general level as that of the English group. The score shown by a child of 13 years 1 month is , by reference to our curve of norms 122 which is considerably higher than the score obtained by these boys. We are,however, unable to verify the above assumption. The results shown by this group of girls is more startling. They appear to be a very inferior group , which . in fact,they are. One of the schools (St Peter's) has among its pupils a large number of orphans . and the general impression gained of them was that they were very retarded and unintelligent. The principal of this school pointed out at the beginning of the testing that the standard of the work done was very low and this she considered was due to the lack of home background , the poor conditions prevailing in their early years and the fact that most of them came from homes that could be considered "poor white". The low median score of the orphan group is only 69 a surprisingly low score.This group (orphans) make up 10% of the total number of girls and must be partly responsible for the lower median scores of the girls.The boys in this group make up only 4% of the total number of boys.

6C. SUMMARY. 1.Norms have been calculated for this area ; all the norms except those of the 9.6 and 14.6 groups are statistically reliable.

2.The Good group of schools have shown a consistent superiority over the girls schools in the Poor group. This is due to the number of children (mostly boys) from upper income group homes who are sent to school in Grahamstown.

3.Boys showed higher median scores than the girls because the Good group constitutes a far larger proportion in their case.

4.Correlation with the South African Group Test was

$+.73 \pm .034$.

5.Correlation with general school work is +.524.

6.The I.Q.'s given by these norms must be multiplied by 1.13 to approximate the results which would be given by the South African Group Test on the same subjects. This factor has little real value as the number of cases from which it is derived is too small. In addition group testing is not accurate enough to establish the I.Q. value of another curve of norms unless the number of cases tested is very large.

7. Afrikaans - speaking children ~~showed~~ achieved scores well below those of the English group. This is due to the fact that a large number of them come from poor homes , a large number of them are orphans and a test in a "foreign" language is not likely to give an accurate picture of ~~the~~ child's ability.

Stastistical Appendix.

1.Means . These were calculated using the following formula with the frequency distribution.

$$M = A + \left(\frac{\sum fd}{N}\right) \times h$$

Where A = the assumed mean ,

f = frequency ,

d =deviation from the assumed mean.(In class intervals),

N =number of cases ,

h = class interval.

2.Medians.

$$Md = 1.1 + \left(\frac{\frac{N}{2} - f_{up}}{f_{md}}\right) h$$

where 1.1 =lower limit of the median interval,

f_{up} = total frequency up to the median interval,

f_{md} = frquency of the median interval,

h = class interval,

N = number of cases.

3. Correlation coefficients were calculated using the product moment formula.

$$r = \frac{\sum xy}{\sqrt{\sum x^2 . \sum y^2}}$$ where x and y are the deviations from the means of each variable.

4.The standard deviation was calculated using the formuha

$$S.D. = \left(\sqrt{\frac{fd^2}{N} - \left(\frac{fd}{N}\right)^2}\right) H$$

5.Skewness. Calculated using the formula

$$Sk = \frac{3\,(M - Md)}{S.D}$$ (Pearson's Approximate Measure of Skewness.)

6. The frequency distributions were drawn up using a classifier recommended by Holzinger for use with small samples. (11) Page 26.

7. The equation for the curve of norms was calculated by the method of averages as described by Holzinger. (11) Page 325 et seq.

8. The probable errors of the means , medians and the standard deviations were calculated from the standard deviation.

List of References. The numbers in brackets before each title refer are those which are used in the text.

BallardP.B. (1)." Group Tests of Intelligence" Hodder and Stoughton .1927.

(2). "Professor Spearman" British Journal of Educational Psychology Vol. XVI Page 1.

Bakes F. (Sister Frances Mary C.R.) "An Investigation into Intelligence Tests used in South Africa with a detailed comparison of the Grey University College Group Scale and the South African Group Intelligence Tests " M.A. Thesis University of South Africa .1930.(3)

Burt C. (4)./"The Factors of the Mind" University of London Press. 1940.

(5). "Critical Notice. The Factorial Analysis of Human Ability by Godfrey Thomson." British Journal of Educational Psychology Vol XVIII Pg 188.

(6) "Methods of Factor Analysis " British Journal of Educational Psychology. Vol.VII Pg 57.

(7). " Mental and Scholastic Tests." P.S.King and Sons. London 1927.

(8). " The Selection of Pupils for Different Types of Secondary Schools." British Journal of Educational Psychology Vol.XVII Page 57.

(9) " Critical Notice. 1.Some Recent Work in Factorial Analysis and a Retrospect.

2.The Factorial Analysis of Human Ability by Godfrey Thomson." British Journal of Psychology Vol XVII Page 40.

Hardie J.L.(10) " Spearman's 'A Measure of Intelligence' , a statistical analysis." British Journal of Psychology .Vol. XIX Pg 188.

Holzinger K.J.(11)."Statistical Methods for Students in Education " Ginn and Co. Boston . 1928.

Holzinger K.J. and Harman H.H. (12). "Factorial Analysis." University of Chicago Press. 1941.

Marshall A.J. (13). " The Standardisation of Professors Spearman's ' A Measure of Intelligence'." Published in "Comparative Intelligence of English,American and Australian children." Melbourne University Press.1934.

Spearman C.(14) "The Nature of Intelligence and the Principles of Cognition ." Macmillan .London .1927.

(15) " The Abilities of Man ." Macmillan . London . 1926.

(16) " The Factorial Analysis of Ability" British Joournal of Psychology Vol.XXX. Page 78.

(17) " A Measure of ' Intelligence ' ; For Use in Schools." Methuen. 1925.

StephensonW(18) " The Factorial Analysis of Ability." British Journal of Psychology Vol. XXX Page 94.

Thomson G.H. (19). " The Factorial Analysis of Ability" ~~British Journal of Psychology~~ University of London Press.1945.

(20). " The Factorial Analysis of Ability."British Journal of Psychology Vol. XXX Pages 105 and 71.

Thurstone L.L. (21)."Nature of Intelligence." Kegan Paul 1924.

(22). " Primary Mental Abilities " University of Chicago Press.1938.

(23) " The Vectors of Mind." University of Chicago Press 1940

Walters E.H. and Thomas F.C. (24). " Some notes on the Standardisation of Professor Spearman's ' A Measure of Intelligence , For Use in Schools." Forum of Education Vol . XIX Page 188.

APPENDIX containing all the score distributions used in the determination of the norms.

1.Table showing the distibution of scores in age groups in the boy's group before the elimination of half the scores in the Good group from the 11.6 group onwards.

Class values.	Age groups.						
	9.6	10.6	11.6	12.6	13.6	14.6	Totals
154.5					2	5	7
144.5		1	8	14	23	35	81
134.5	1	10	20	20	45	68	154
124.5	4	9	17	31	29	33	123
114.5	5	15	27	19	27	17	110
104.5	3	15	20	7	3	8	61
94.5	11	19	14	10	8	5	67
84.5	3	14	4	4	2	1	28
74.5	10	11	3	3	1		28
64.5	2	3	1	1			7
54.5	2	1					3
44.5	1						1
Totals.	41	89	114	109	145	172	670.

The totals in this table are given in Tables IV and V , pages 35 and 37 respectively.

2. Table showing the distribution of scores in age groups in the girls'group.

	Age groups						
Class values	9.6	10.6	11.6	12.6	13.6	14.6	Totals
154.5							
144.5			2	4	9	14	29
134.5		1	2	15	22	24	60
124.5	1	7	10	17	18	17	70
114.5	3	8	12	19	15	13	70
104.5	11	13	13	10	7	12	66
94.5	7	11	11	7	2	3	41
84.5	6	6	9	4	1	1	27
74.5	5	4	7	4	3	1	24
64.5	3	2	4				9
54.5	1	1		1			3
Totals	37	53	70	77	77	85	399.

The totals in this table are given in Tables IV andV , Pages 35 and 37 respectively.

3. Table showing the distribution of scores in the Good group before the elimination of half the scores from the 11.6 group onwards.

	Age groups.						
Class values.	9.6	10.6	11.6	12.6	13.6	14.6	Totals.
154.5					2	4	6
144.5			6	14	19	26	65
134.5	1		16	17	38	57	129
124.5	2	7	15	28	25	24	101
114.5	3	15	19	12	13	12	74
104.5	1	12	14	3	4	6	40
94.5	9	8	7		2	1	27
84.5	2	8	1	1			12
74.5	4	3	1	1			9
64.5	1			1			2
54.5	2						2
Totals	25	53	79	77	103	130	467

The totals in this table are shown in Tables IX and X , both on page 41.

4. This table shows the distribution of scores in the Poor group.

	Age groups.						
Class values.	9.6	10.6	11.6	12.6	13.6	14.6	Totals.
154.5						1	1
144.5		1	2		4	9	16
134.5		4	4	3	7	11	25
124.5	2	2	2	3	4	9	22
114.5	2		8	7	14	5	36
104.5	2	3	6	4	4	2	21
94.5	2	11	7	10	6	4	40
84.5	1	6	3	3	2	1	16
74.5	6	8	2	2	1		19
64.5	1	3	1				5
54.5		1					1
44.5		1					1
Totals	16	36	35	32	42	42	203

The totals in this table are shown Tables IX and X, both on page 41.

5. Table showing the distribution of scores in age groups in the Good group after the elimination of half the scores in this group from the 11.6 group onwards.

666

Class values	Age groups 9.6	10.6	11.6	12.6	13.6	14.6	Totals.
154.5					2	1	3
144.5			4	6	9	15	34
134.5	1		7	10	19	32	69
124.5	2	7	9	14	18	8	52
114.5	3	15	6	6	6	4	40
104.5	1	12	7		3	4	27
94.5	9	8	4			1	22
84.5	2	8	1	1			12
74.5	4	3	1	1			9
64.5	1						1
54.5	2						2
Totals	25	53	39	38	51	65	271

6.Table showing the distributionof scores in age groups in the Good group. Girls.

	Age groups						
Class values.	9.6	10.6	11.6	12.6	13.6	14.6	Totals.
154.5							
144.5			1	3	5	4	13
134.5			1	6	10	4	21
124.5		3	3	3	4	5	18
114.5	1	2	3	3	3	2	14
104.5	4	3	1	2		1	11
94.5	3	4	1	1			9
84.5		1	3				4
74.5							
64.5			1				1
Totals	8	13	14	18	22	16	91

The totals in this table are shown in Tables XIII and XIV on pages 44 and 45 respectively.

7. Table showing the distribution of scores in age groups in the Poor group.Girls.

Class values.	Age groups. 9.6	10.6	11.6	12.6	13.6	14.6	Totals.
154.5							
144.5			1	1	4	10	16
134.5		1	1	5	12	20	39
124.5	1	4	7	14	14	12	52
114.5	2	6	9	16	12	11	56
104.5	7	10	12	8	7	11	55
94.5	4	7	10	6	2	3	32
84.5	6	5	6	4	1	1	23
74.5	5	4	7	4	3	1	24
64.5	3	2	3				8
54.5	1	1		1			3
Totals	29	40	56	59	55	69	308

The totals in this table are shown Tables XIII and XIV on pages 44 and 45 respectivaly.