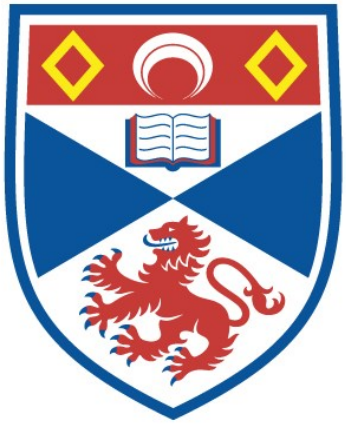
AN INVESTIGATION OF METHODS OF EXAMINATION AT THE 'QUALIFYING' STAGE

Douglas Moul McIntosh

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



1933

AN INVESTIGATION OF METHODS OF EXAMINATION AT THE "QUALIFYING" STAGE

being a Thesis presented by

DOUGLAS MOUL McINTOSH, M.A., B.Sc., B.Ed.

to the University of St. Andrews

in application for the degree of Ph.D.

CERTIFICATE.

I certify that Douglas Moul McIntosh has spent nine terms at Research Work carried through chiefly in Dundee, that he has fulfilled the conditions of Ordinance No. 16 (St. Andrews), and that he is qualified to submit the accompanying Thesis in application for the Degree of Ph.D.

.....

30th January, 1939.

DECLARATION

I hereby declare that the following Thesis is based on research carried out by me, that the Thesis is my own composition and that it has not previously been presented for a Higher Degree.

The main locus of the research was Dundee.

CAREER

I matriculated in the University of St. Andrews in 1927 and graduated in June 1931 with the degrees M.A. and B.Sc., with first class honours in Mathematics. In 1932 I obtained the Diploma in Education of the University of St. Andrews. I matriculated in the University of Edinburgh in 1934 and graduated with the degree of B.Ed. in June 1935.

In October 1935 I commenced the research on methods of examination at the "Qualifying" stage which forms the basis of the Thesis now being submitted in application for the degree of Ph.D.

AN INVESTIGATION OF METHODS OF EXAMINATION AT THE "QUALIFYING" STAGE.

AN INVESTIGATION OF METHODS OF EXAMINATION AT THE "QUALIFYING" STAGE.

PART I.

PART II.

PART I.

I. INTRODUCTION

From earliest times, examinations have played an important part in the systems of education adopted by communities, barbaric and civilised. The youths of the primitive tribe were subjected to examinations in the form of initiation ceremonies in which they had to achieve certain physical and mental standards before being admitted as adult members of the tribe. This practice was a feature of Greek education, particularly among the Spartans, from about 500 B.C.[1] There are, also, records dating back three or four thousand years, of a national examination system in Ancient Chinese civilisation which had as its object the selection of public officials.[2]

1. The Schools of Hellas: Kenneth Freeman. Macmillan & Co.Ltd.1907

2. The Meaning of National Examination System, Pan. Q. China Critic, 1928, 1, 512-515.

Examinations of Western educational systems have as their prototypes those of the mediaeval universities. The analogy has been drawn between the early universities and the trade and craft gilds of the same period, the former in fact being teachers' gilds. Just as trade gilds had their apprentices, journeymen and masters, so the universities had their scholars, bachelors and doctors. As in the case of the gilds/

gilds, the various stages could only be completed when the candidate had performed a piece of work to the satisfaction of the gild, so each stage in the university could only be successfully accomplished by passing an examination set by the University Authorities. The scholar had to undergo a test of his ability to define and explain terms, success in which entitled him to his baccalaureate degree and, on concluding his bachelor's course, before he became a doctor or master he was compelled to teach in public or to defend a thesis before the University Authorities.

Two points in connection with the examinations of the Mediaeval Universities may be observed. Firstly, the examinations were oral; secondly, they fulfilled the purpose for which they were intended, namely, to find whether the candidates could accomplish a task similar to that with which they would be confronted in their post graduate situations.

Written examinations seem to have been a product of the early 18th Century. In 1702, Cambridge University introduced the written examination, a system which was adopted soon afterwards by Oxford University and the other centres of learning and has been continued down to the present day. With the extension of the function of the Universities the oral examination with its essentially practical character gradually fell into decay and the written examination assumed other functions which/

which affected considerably not only the examinations but the system of education itself.

Dissatisfaction with the examination system began in England and America about the year 1880. Professor Edgeworth of Oxford University published practically the first important statistical investigation on examinations in three papers[1]

1.
(i) The Statistics of Examinations, Journal of the Royal Statistical Society, vol. 51, pp 599 - 635 (1888).

(ii) The Element of Chance in Competitive Examinations, Journal of the Royal Statistical Society, vol. 53, pp 460-75 and 644-63. (1890)

(iii) On Problems in Probabilities, Philosophical Magazine, August 1890.

---

on the results of examinations.

In these papers he drew attention to two problems, firstly, the discrepancies which might arise when a number of different examiners marked the same examination scripts, and secondly, when the same examiner marked the same scripts a second time at a later period.

With regard to the constancy of marking Professor Edgeworth found that in the First Class Civil Service Examinations the element of chance was so great that only one third to two thirds of the successful candidates could be sure of success if a different set of equally competent examiners had been appointed. In another experiment, a piece of Latin Prose when marked/

marked by twenty eight highly qualified examiners was awarded marks varying from 45 to 100 out of a possible mark of 100.

Dealing with the second problem, he found that in a typical examination the total probable error due to such errors as correction at great speed, fatigue, and the examiner's idiosyncracies, was more than two per cent of the aggregate of marks. "For instance, if a candidate obtains an aggregate of 1,965 marks, thus falling below the Honour line by 35 marks, the odds that he ought by right to be in the Honours Class are considerable, say one in four. Even if he has obtained only 1,800 marks, it is by no means certain that he is rightly placed among the rank and file of Pass Candidates."

At about the same time, in America, several of the prominent educationists were protesting against the system of examinations which was rapidly dominating education. President Charles W. Eliot of Harvard College attacked the examinations from the point of view that they treated all pupils alike, ignoring the individual differences of the children.[1] His

---

1. Educational Reform. C.W. Eliot, New York. 1901.

---

plea was for the adaptation of the examinations to individual capacities, though at this time the technique of measurement had not developed far enough to accomplish such a task, and the only substitute he could suggest was teachers' estimates.

An/

An attack from a different angle was made by Professor J. Rendell Harris of Haverford College, whose claim was that the magnitude of the mark mattered little in comparison with its position in relation to the others of the group. "A well conducted examination divides the students one from another like the opening of a fan. I affirm that the first thing to be aimed at is to produce a dispersion among the group of persons presented for examinations."[2.]

---

2. Proceedings of the First Annual Convention of the College Association of the Middle States and Maryland, 1890.

Indeed, what Professor Harris was advocating was that the marks of an examination should be distributed along a normal frequency curve, a technique which was to be introduced some twenty years later with the development of statistical research.

With the rapidly changing philosophy and outlook of the early twentieth century and the development of the new techniques in educational and psychological measurement, the question of the place and functions of examinations in education was being critically examined in practically every country. When the New Education Fellowship met at Locarno in 1927, an international commission was appointed to look into the subject. In 1929, at the Elsinore Conference, there was a discussion of the evidence submitted by the representatives of some twenty countries. That the subject of examinations should be studied internationally was felt by all representatives and it was decided/

decided to continue the investigations. After a Conference which was held at Nice in 1932 a report, "Examinations and the Way Out", was published in 1935.

The next step in the organisation of international resources for the attack upon examinations came with the establishment of the International Examinations Inquiry, under the sponsorship of the Carnegie Corporation, the general supervision of the work being undertaken by Professor Paul Monroe, Director of the International Institute of Teachers College, Columbia University. The first Conference was held at Eastbourne in 1931 when twenty eight representatives from England, France, Germany, Scotland, Switzerland and the United States discussed the problems connected with examinations, a report of the Conference being published under the title "Conference on Examinations"[1].

---

1. Bureau of Publications, Teachers College, Columbia University. 1931.

In his opening address, Dr. Monroe pointed out that, in addition to the problems concerning examinations themselves, there was a great need for a study of the effect of examinations upon the social structure of the nation. Scant attention, he declared, had been paid to examinations as a means of social control. In most of the European countries, for example, the gateway to all the professions and government service was guarded by examinations which therefore formed the foundation of/

of the intellectual class. America, on the other hand, did not have the same intellectual and cultural tradition on which to base a system of education and consequently an intellectual class did not exist. Another problem, however, had arisen, namely, the tendency to produce a greater number of those prepared for intellectual, professional and clerical activities than could be well absorbed by the respective societies.

At the Eastbourne Conference there was a general consensus of opinion that in each country the problem of examinations had reached a different stage of development, and the only method which could be adopted with any success was for each country to undertake an investigation appropriate to their own situation. It was decided, therefore, that national committees should be set up with this aim in view. In England and France, for example, the unreliability of the marking of examinations had not been appreciated to the same extent as in Scotland and America, with the result that the committees in the former countries undertook inquiries on this point before considering any further problems.

In Scotland two investigations by the Scottish Council for Research in Education had already been started dealing with teachers' estimates and the time devoted to examinations. These were completed and published in 1932 and two new investigations were undertaken. One, a mental survey of Scottish children/

children and a subsequent follow up of a selected group, and the other dealing with the predictive value of the Leaving Certificate Examination. In Germany as in Scotland, a committee had already been at work on the problem of examinations, their particular subject being the selection of children for admission to the secondary schools. This investigation was continued and another was undertaken by Dr. Robert Ulich and Dr. Erich Wohlfahrt, the aim of which was to show the relations between the social origin of students, the type of secondary school which they selected, their professional aims, and the success in the courses undertaken. The Swiss representative, Professor Bovet, undertook the task of preparing the history of the examination of army conscripts in Switzerland.

In June 1936, at the third Conference on Examinations, the final reports of the investigations were presented, an Interim Conference having been held in London in 1933. The work of the English committee had been published under "An English Bibliography of Examinations (1900-1932)", "An Examination of Examinations", "Essays on Examinations", and "The Marks of Examiners". The French investigation which deals mainly with the Baccalaureat, the examination at the close of the secondary school, is entitled "La Correction des Epreuves Escrites dans les Examens". In Germany, O. Bobertag has published the investigations of the committee under the title "Schulerauslese"/

"Schulerauslese" and in America, Professor Kandel has prepared a report entitled "Examinations and their Substitutes in the United States." Two volumes have been published by the Scottish Committee, namely, "The Intelligence of Scottish Children : A National Survey of an Age Group" and "The Prognostic Value of University Entrance Examinations in Scotland." Another investigation, also, has been started by the Scottish Committee dealing with the methods of examination at the "qualifying stage" and it is some of the results of this investigation with which the second part of this thesis is concerned.

II. HISTORY OF QUALIFYING EXAMINATION

In 1901, the Education (Scotland) Act raised the age of compulsory attendance at school from thirteen to fourteen and thus some instruction of a more advanced nature than that given in the elementary school was possible for the majority of the pupils. A re-organisation of the school system was therefore necessary. Provision for two years courses of specialised instruction, called supplementary courses, was made in 1903 by the Code of that year.

Entry to such courses was determined by a pupil's success in an examination - the qualifying examination. Although the establishment of the qualifying examination is of fairly recent origin, the idea of providing a 'hurdle' between elementary and secondary education may be traced back to the time of the Reformation.

In his First Book of Discipline, Knox devotes a chapter entitled, "Schools and Universities", to education where he sets out a complete national system of education with a corresponding national system of examinations. For the purpose of examining the pupils at the end of each grade it was decreed that "discreet, learned, and grave men be appointed to visit all schools for the trial of their exercise, profit and continuance; to wit the ministers and elders, with the most learned men in every town, shall every quarter take examination." The/

The examinations like all others at this time were oral in character.

These visitations of examiners became a feature of school life and the day on which the visitation took place was an event both to scholars and teachers. Teachers were anxious that their pupils would show up as well as those of neighbouring schools and thus efficiency of instruction and discipline were maintained. As the examination was carried out with great thoroughness, a certain uniformity of instruction was kept throughout the schools. Despite many efforts to deprive the church of the right of inspection and visitation, this privilege was retained till the Education Act of 1872.

It was evident by the middle of the nineteenth century that some national body would have to be created to supervise the distribution of the grants for education made by the government. Accordingly, in 1872, by the Education Act of that year, there was established the Scotch Education Department. This body included among its functions the annual inspection of schools. By the Act financial provision was made for the Parish Schools which were not limited to elementary education, the School Boards being enjoined to conserve the secondary education given in these schools. The Burgh Schools were non "state aided" and were known as Higher Class Public Schools under School Boards. The demand for secondary education/

education steadily increased and provision had to be made either by the development of the secondary education provided in the Parish Schools or by increasing the number of Burgh Schools which would then be brought into the national system as providing secondary (as distinct from primary) education. It was the former method which was chosen and thus a national system was evolved, in keeping with the ideal of John Knox to establish distinct grades of national education leading from elementary through secondary education to the Universities.

Several factors assisted the development of secondary education within the state-aided schools and the steady merging of the Burgh Schools into this system. Among these were, the introduction in 1899 of Higher Grade Schools providing a three years' post primary course, many of which developed into full secondary schools; the development of two years' post-primary courses known by the name of Supplementary Courses being of four types - Commercial, Industrial, Rural and Domestic; the establishing of the Leaving Certificate Examination in 1888 to mark the successful completion of a full secondary course. These, along with a careful manipulation of grant regulations by the Department assisted in bringing the Burgh Schools into line, but what was perhaps the most important factor of all was the demarcation between primary and secondary education at about the age of 12.

At/

At this point, brief mention may be made of an examination system introduced into the schools during the nineteenth century, namely, the system of payment by results. In short, the method was that the state grant was dependent on the examination success of the pupils as determined by the inspectors' reports. Fortunately such a system was not tolerated for any length of time in Scotland though its effects were felt for many years after the system had been abolished. In 1885 all traces of the system were removed from the lower standards and by 1890 it was abolished completely from the elementary school.

Elementary education became practically universal by the Education Act of 1872. To give a greater definiteness to its aim, the merit certificate was introduced by the Education Department, thus providing a goal for the elementary school child similar in nature to that afforded by the Leaving Certificate for the secondary school pupil of to-day.

The Code of 1892 laid down the necessary requirements for the attainment of this certificate. Briefly, these were that the child should be over thirteen years of age; that he should be thoroughly proficient in the three R's; that he should be proficient also in two class subjects such as English, Geography, History, Needlework, and Elementary Science; and that he had passed satisfactorily through all the grades of one specific/

specific subject. The prime motive in this was that each child should receive a thorough grounding in the three R's, although it was hoped that many would be encouraged to remain at school longer than the compulsory school leaving age.

Up to this point most of the legislation and movements for reform had been directed towards the elementary school, but it soon became evident that the secondary school system was also in need of similar attention. In 1898, the Scottish Education Department carried out reforms with this end in view. The merit certificate was now given the function, not of marking the completion of the elementary school as previously had been the case, but to provide a line of demarcation between primary and post-primary education. It acted as a hall-mark of the satisfactory completion of a primary course, and a guarantee of fitness for entry upon a post-primary course. To this end, the age qualification was reduced to twelve and the regulations regarding specific subjects withdrawn. In the words of a Scottish Education Department Memorandum, what was required of the pupil was that he should have "the ability to read, write, speak and understand plain English and to perform simple calculations."

The actual requirements are given in Article 29 of the Code. A pupil is expected

"(a) To read at sight, with good pronunciation and with intelligent/

intelligent phrasing, narrative prose of moderate difficulty.

(b) To write to dictation, with good spelling and legible and regular handwriting, a narrative passage previously unseen.

(c) To answer questions as to the subject-matter of and the meaning of words and sentences in the reading books in use in the class; these answers, when necessary, to be expressed in complete sentences or in a consecution of sentences.

(d) To write a composition, the heads being given, or to give in writing, the substance of a passage read.

(e) To know the four rules of Arithmetic as applied to whole numbers, easy vulgar fractions, and decimals to three places, and to be expert in applying this knowledge to the calculation, both mentally and on paper, of simple sums in money and in common weights and measures.

(f) To be reasonably proficient in the other subjects included in the approved scheme of work of the class."

It is to be noted also that this examination was not conducted by Local but by Central Authority and accordingly was uniform throughout the country, apart from individual differences among the inspectors.

Still another change necessitated by the raising of the age for compulsory attendance at school from thirteen to fourteen by the Education (Scotland) Act of 1901 was introduced in/

in 1903 by the Code of that year. It was felt that an advanced course of instruction should be available to those who would be leaving at the age of fourteen and who had successfully completed their elementary course at the age of twelve. Accordingly, two years "supplementary courses" were introduced. To mark the completion of these new courses, the merit certificat was transferred to this stage and its place was taken by a qualifying examination which carried with it no certificate.

The next stage in the history of the Qualifying Examination was reached in 1921 when by Circular 44 the Department intimated the abolition of the Qualifying Examination as conducted by H.M. Inspectors. The importance of the break between primary and post-primary education was not forgotten and it was pointed out that the qualifying stage must "mark something of an epoch in the school career of every pupil." The various education authorities were advised to come to "some arrangement under which those who have taught and those who are to teach particular individuals shall combine in an endeavour to estimate the potentialities of the material to be handled."

As to the method the Department suggested

> "It is neither necessary nor desirable that any rigid uniformity of method should be insisted upon. What is practicable and convenient in one education area may conceivably/

conceivably be unsuitable in another. Broadly speaking, however, the Department will expect that the standard of the existing Qualifying Examination, the soundness of which has been proved by an experience of nearly twenty years, will be reasonably maintained. That is, any test to be applied should be such as the average pupil of 12 may fairly be called upon to meet."[1.]

---

1. Scottish Education Department Circular 44, par. 2 and 3, 12th December, 1921.

An account of the methods adopted by the various Authorities is given by Professor McClelland in the 1934 Year Book where he gives a summary of the results of a questionnaire to the education authorities. Slightly less than half retain a uniform written examination, the subjects being English and Arithmetic, with the addition, in a proportion of cases, of history and geography. "Sometimes the papers are set by an examination board, sometimes by the Director of Education, external examiners or head teachers. Even in these areas considerable weight is given to the school records of the pupils, and in cases where the uniform examination has been abandoned these are the main basis of the decision. Formal oral examination is uncommon, but twelve areas are making use of intelligence tests, and five areas of standardised scholastic tests.

While/

While in about fourteen areas the passing of the qualifying test is all that is demanded for entrance upon a secondary course, the majority of authorities place a higher hurdle before the door of these courses. In eight areas the same test is used for this purpose, but a higher mark is required. In eight other areas there is a special examination - usually called a "control" examination - for entry to the secondary departments; and in a number of cases the secondary schools have their own entrance tests. As a rule, consultation between the receiving and transmitting headmasters is provided for, either in all or in doubtful cases."

Such was the position in 1934 but there has been considerable change since then and one of the functions of the International Examination Enquiry Committee's Qualifying Investigation will be to provide a statement of the position as it is to-day. One of the greatest changes which has taken place is the change in <u>attitude</u> towards the Qualifying Examination. Formerly it was regarded as a barrier or hurdle which the child must clear before passing into a post-primary school, but now the great range of individual differences in children is being recognised and the Qualifying Examination, no matter what its form may be, is assuming a prognostic function.

Formerly, too, examinations looked towards the past. A candidate gained admission to a secondary school by passing an/

an entrance examination. This method of selection, however, gave no guarantee of success in whatever course he followed in that school. To-day the tendency is to look forward to the future. An attempt is being made to frame the tests so that they will give some reliable prediction of a candidate's likelihood of success in the various post-primary courses of instruction. For this purpose it is not sufficient to find out what he already knows.

Little is known as to the prognostic value of the various types of tests and examinations used by authorities at the Qualifying stage. It is with this problem that the following investigation is concerned.

It has been seen that one of the main points in the development of secondary education was the establishment of a clear line of division between primary and secondary education. The transfer from the one to the other generally took place at about the age of 12 years, although the less able children unable to surmount the examination barrier between the two types of education were much older. One of the most significant trends in the schemes of promotion to post-primary education is the "clean cut" promotion of pupils at about 12 years of age, especially of the pupils unable to qualify on the basis of achievement.

The "clean cut" grew out of the recommendation of the Department/

Department that the now qualifiable type of child should be promoted on the basis of age alone not later than the age of 13 years.[1] The Department Circular 44(1921) stated "there need

1. Scottish Education Department, Circular No. 44, Dec. 13, 1921, page 2.

be no restriction as to the age at which promotion takes place. If a pupil is ready to move forward before he is 12 years old, and if he can do so without risk of over-pressure, no hindrance should be placed in his way. Nor will it be advisable to require that those who are below the average in ordinary school subjects shall necessarily remain in the "qualifying" class until the obstacle has been surmounted. In the larger schools, at least, there will always be a not inconsiderable proportion whose tastes are anything but bookish, and whose interest can be more effectively awakened through the hand than through the head ..... To require that boys and girls shall spend two or three years of their lives in striving after the unattainable is as futile from the point of view of the State as it is cruel from the point of view of the individual."

Again, there was the growing recognition of the value of practical post-primary courses and the need for a longer period of time to devote to such courses in the advanced division schools.[2] The Primary School Code in 1923 and the Sub-

2. Code of Regulations for Day Schools in Scotland, Art.15, 1923.

Committee of the Department's Advisory Council, 1923, urged[1]

---

1. Report of the Advisory Council to the Scottish Education Department on the General Organisation of Day Schools and of Continuation Schools and Classes, 1923, pp. 3-4.

---

that pupils who were unable to qualify by achievement when they had reached the age of 13 years should be automatically promoted to a post-primary course where a special curriculum might be developed for them. The 1923 Code reads "As a rule, no child should be promoted to a higher division who has not satisfactorily completed the work of the division below, but education authorities in all cases may, and in the case of children over 13 should, make special provision for the instruction of children who in the course of normal promotion have failed by more than a year to reach the division appropriate to their age."[2] The Committee on Education and

---

2. Committee on Education and Industry, Second Report, 1928, p.17.

Industry in Scotland in 1928, were strongly in favour of the principle of the "clean-cut" at 13.

It is to be remembered that the Hadow Committee in their report on "The Education of the Adolescent in 1926" supported the tenet that there should be automatic promotion at the age of 11+ to a suitable course of post-primary instruction. Scotland, therefore, had been slowly adopting such a system for/

for several years before the publication of the Report.

As to the actual position in Scotland it would appear as if most of the authorities are adopting the "clean-cut" system at 12 years of age, although only a minority have qualified at this age. As a result of a questionnaire in 1934, Professor McClelland found that 60% of the authorities reported 12 to be the normal age of transfer. Some would prefer a reduction to 11 plus, while very few favour 11 or 12 plus.

III. SURVEY OF PREVIOUS FOLLOW-UP INVESTIGATIONS AT THE QUALIFYING STAGE

A. INTRODUCTORY

In his contribution to "Essays on Examinations",[1] Spearman distinguishes between reliability and validity. The former is the degree of agreement between any two independent sets of measurements of the same set of things while the latter is the agreement of measurements with the things measured. He points out that of the two validity is by far the more important, but unfortunately this aspect of examinations has been, to some extent, neglected, largely due to difficulties involved in any investigation dealing with such a problem. The most promising method of dealing with the validity of examinations is that of tracing the subsequent careers of the examinees. Spearman concludes his article with the hope that future research on the subject of examinations would extend its scope so as to include their "follow-up".

1. Essays on Examinations:- Note on the Reliability and Validity of Measurements by C. Spearman, International Institute Examinations Enquiry. Macmillan, 1936. p. 108.

"Follow-up" investigations have been comparatively few though several experiments were carried out in America, particularly with students entering the Universities. The Scottish Committee of the International Institute took up this question/

question, also at a University stage, and the results of their investigation have been published in "Prognostic Value of University Examinations in Scotland", (University of London Press, 1936).

Little has been done to test the validity of the various methods of selection of pupils particularly for post-primary courses of instruction.

B. BRITISH INVESTIGATIONS

Prior to 1924 investigations on examinations had been concerned mainly with reliability and the introduction of intelligence tests as one of the methods of selecting pupils for secondary education. An account of some of the experiments conducted in the use of tests in free place examinations in England is given in the Government report on Psychological Tests of Educable Capacity.[1]

1. Psychological Tests of Educable Capacity. Board of Education Report, H.M. Stationery Office, 1924. Appendix II, p. 149.

Several areas had been introducing and experimenting with intelligence tests in their methods of selecting children for secondary schools. It was mainly due to discrepancies in the results of these tests and the results of "control" examinations in English and Arithmetic which gave rise to further experiment on the question of the validity of these examinations. Actually/

Actually there seems to have been no large investigation. Most of the reports were of a general nature citing particular cases in support of specific arguments.

In the County Borough of Blackpool, experience with tests seems to have indicated that the test results were a helpful supplement to the usual scholastic examination. Although no definite experiment is quoted, the Director states "that a comparison of the work of the pupils, subsequent to the application of the tests, was showing that these tests did assist in the discovery of educable capacity."

In 1922, an intelligence test was used in Bradford as well as a scholarship examination for the awarding of free places to secondary and central schools, although the test results were used only with border line cases. When pupils were awarded scholarships mainly due to their score in the intelligence test, it was found that these pupils generally justified their selection.

An investigation with 48 boys and 44 girls revealed that, in general, the results obtained from the intelligence tests corresponded with the progress made in the secondary school, as indicated by the terminal examinations, to a rather greater extent than obtained when a comparison was instituted between the marks awarded in the ordinary scholarship examination and the marks given in the secondary schools.

In/

In the County of Northumberland more, perhaps, has been done with intelligence tests than in any other area. Professor Godfrey H. Thomson, who was largely responsible for much of the early work, reported that free places were given to any who scored highly in either the examination or the test provided they were average in other respects. In 1921, however, fourteen places were given on the result of an intelligence test alone, and of these, only one did unsatisfactory work in the secondary school.

The Director of Education for Northumberland has also submitted statistics which indicate that pupils who had low marks in the examination but intelligence ratios of 120 or more were all doing satisfactory work in the secondary school with the exception of one pupil. On the other hand, cases are quoted of pupils who obtained high marks in the examination but low marks in the test and whose progress in the secondary school was unsatisfactory. His general conclusion was that "the introductim of a psychological test, as a supplement to the examination in school subjects, has been of great value".

Another interesting experiment was that carried out by T. P. Tomlinson, Head Master of Love Lane Council School, Pontefract, who tested with a group intelligence test 1600 children distributed among 16 schools in the West Riding of Yorkshire./

Yorkshire. The test had a reliability of ·95 and correlated with individual Binet testing to the extent of ·71 to ·84. From the data it is concluded that the intelligence test reveals ability not shown by the examination due to such causes as nervousness, home circumstances and the like. Also, that success in examinations may be due not so much to ability as to diligence and favourable home conditions. The intelligence test often reveals a low mental level thus hidden and the subsequent school career indicates the pupil's inability to cope with the increasing difficulties of the curriculum. In one central school 2% of the children had I.Q.s. below 85, their success in the entrance examination being probably due to special preparation. The head master's report on the work of these pupils indicated that they were unable to make a success of the course to which they had been promoted.

Similar conclusions were reached by R. R. Dobson, Head Master of Pate's Grammar School, Cheltenham, in experiments conducted by him in the selection of pupils for free places to secondary schools. This report is of interest from the point of view that the investigator, in one experiment, used an oral intelligence test (Burt's graded reasoning tests). In his selection of 26 from some 150 candidates he awarded some free places solely on the grounds of the pupils' success in the intelligence test. The reason given was that these pupils

came from schools in which the level of teaching was inferior to that in the normal city schools. The subsequent progress of the pupils showed that the selection had been fully justified. Another important point is made in the report, namely, that some sort of an examination is required as a certain level of attainment is necessary before a child can take advantage of a secondary school course.

In a published report of this experiment,[1] the author

---

1. The Value of Intelligence Tests in Scholarships Examinations. R.R. Dobson. The Forum of Education, Vol, I, No. I, 1923.

---

defined the object of an entrance examination to be the selection of "the most mentally alert boys as distinct from those whose standard of attainment is high in virtue of efficient teaching, cultured homes and even careful coaching for the examination itself." He pointed out, also, that the difficulty is complicated by the fact that the quality of the teaching staff varies very much from one elementary school to another, and it would appear that the teaching in elementary schools in England varied far more than was generally supposed.

The correlation between the reasoning tests and the written examinations was ·41 and this coefficient, when worked out for the school which sent the greatest number of candidates, becomes ·56. This emphasises the fact that where teaching is uniform, for example, in the same school, whether it be good or bad/

bad, there the written scholastic examination may give a fair estimate of comparative intelligence among the boys of that school; but where the teaching conditions differ greatly, there such an examination by itself is less just.

In an investigation on intelligence tests[1], Oates

1. A Statistical and Psychological Investigation of Intelligence Tests. D.W. Oates. The Forum of Education. Vol. VI, No.I, 1928. p. 38.

came to the important conclusion that much greater care was necessary in refining tests to be used in discriminating between the relative intelligence of boys of average ability than in devising tests for selection of the brightest boys.

It is worth emphasising that this conclusion holds not only for intelligence tests, but selective examinations of all kinds. It is usually a fairly easy task to differentiate between the outstanding pupils as also is the case with the very dull children. The real difficulty arises among borderline cases and, as one mark often decides whether or not a pupil is to gain a free place to a secondary school, it is essential that the examination must have a fine discrimination at this point. The following two investigations give weight to this argument.

Oates, in an investigation, tested with Ballard's Chelsea Test, Columbian Test and Crichton Test, 270 secondary school boys whose ages ranged from 11 to 18 and who were distributed/

tributed among ten forms. For each pupil, besides the ordinary school examination marks, there were prepared independent rank estimates for intelligence by the two masters who had the best knowledge of the boys in each form. It is interesting to note that the coefficients of reliability of the estimates were ·69, ·70, ·83, ·48, ·69, ·67, ·77, ·26, ·72, ·85 for the respective forms. The two separate estimates were given equal weights and combined giving a figure which was used as the teacher's estimate of intelligence. These estimates of intelligence seemed to be influenced, to some extent, by success in school examinations, a fact which was shown in another investigation.[1.]

---

1. The Nature and Validity of Subjective Estimates of Intelligence Forum of Education, II, 1924. pp. 103-121.

In the case of parallel forms A and B, in ten out of the eleven cases, teachers' estimates yielded a higher coefficient with intelligence in the superior form than they did in the poorer form "which suggests that there is some constant factor operating in these cases." Thus it is concluded that differentiation is easier among pupils of superior ability than among pupils who approach more nearly the median for the normal distribution of ability.

This/

This same point was raised by Valentine in his investigation on the Reliability of Examinations[1] where he

1. The Reliability of Examinations - An Enquiry. C.W. Valentine. University of London Press, 1932. Chap. III. pp.43 - 50.

reveals the weakness of all classification, namely, that children cannot be divided simply into those fit and those unfit for secondary education. There is always a large borderline group of whom it is extremely difficult to say whether they fall in either the one category or the other, "consequently a delicacy of discrimination in the examination, especially near the borderline of passes, is extremely important."

Valentine, in a preliminary discussion of his investigation, shows how in one centre, at which the entrance examination is considered to be fairly reliable, 130 children lie within about 1%. of the admission mark and of these only 65 have to be selected. "If nothing further is done, one might almost as well put the names in a bag and draw out 65,", is his comment. The subsequent careers of these children also indicates the seriousness of the problem. Out of 36 pupils who constituted the last three in twelve entrance examination lists in this centre, 17 gained a School Certificate four or five years later. Now, considering the pupils whose entrance/

entrance examination consisted of (1) intelligence tests, (2) written papers in English and Arithmetic, (3) an oral test, the candidates with 60% or over gained admission. There were 18 successful candidates, of whom 12 were only separated by 3%. The last four in the list, all of whom scored the borderline 60%, were the most distinguished pupils of the group in their secondary school course. Every one gained a School Certificate, one with matriculation, while the top boy in the entrance examination failed to obtain a School Certificate. That this is a general result is shown by the fact that, from the results of five centres, 30 out of 67 children who were in the bottom third at entrance, gained a School Certificate.

That success in the secondary school is not necessarily measured by success in gaining a School Certificate is admitted but the conclusions are also proved when progress in the secondary school is measured by the Head Teacher's reports. In six large schools no less than 140 out of 797 left school with bad reports in the first four years. The following shows the position of these pupils in the entrance examination:-

Top	third	at	entrance	...	14·3%	had	bad	reports
Middle	"	"	"	...	19·1%	"	"	"
Bottom	"	"	"	...	19·2%	"	"	"

"Thus in entrance examinations even better than the average of those/

those tested, the examination order fails to discriminate clearly between the good and the bad, and fails to exclude many who are apparently unfit for secondary education, taking as our criterion the report of the Heads. Here again then we see how important it is that the order of merit at the entrance examination should be a reliable one."

Amos carried out an experiment in an industrial area with 2000 children whose ages ranged from $10\frac{1}{2}$ to $14\frac{1}{2}$ years with a median age of $12\frac{1}{4}$ years, sitting a secondary school entrance examination.[1] One third of the candidates gained

---

1. Examination and Intelligence Test Forecasts of School Achievement. A. Donald Amos. Brit. J. Educ. Psych., 1931, Vol. I, Part I, p. 73.

---

places in secondary schools and the follow-up involved 270 pupils from different schools. The results deal mainly with the first two years in the secondary school. Raw scores are used in the calculations and age allowance is not dealt with "arising from the consideration that we were engaged in the investigation of a problem under conditions presented by the existing system."

The first year results given in Table I (page 34) show that the correlations with school marks are higher in the case of the intelligence test than with the entrance examination while a combination of the intelligence test and the examination/

examination gives a higher correlation than either the test or the examination. From these first year results the investigator concludes that "we have strong grounds for recommending the inclusion of intelligence tests in entrance examinations for admission to secondary schools. If we have to chose one or the other, then the intelligence tests are preferable to the usual Arithmetic - English examination, but the strongest arguments of all advocate the inclusion of both types of examination."

TABLE I.

Group	n	r_{ac}	r_{bc}	r_{qc}
Mq	30	.43	.50	.53
Km	30	(.19)	.26	.34
Ea	31	.31	.32	.37
Hk	32	(.26)	.35	.43
B	33	.47	.44	.57
C	48	.29	.41	.40
D	70	.35	.49	.54

a = Forecast by Entrance Examination in English and Arithmetic.

b = Forecast by Intelligence Test

q = Forecast by Entrance Examination and intelligence tests combined.

c = Actual achievement at the end of the first secondary school year.

(The coefficients in brackets have no significance)

TABLE II/

TABLE II.

Group	n	r_{ac}	r_{bc}	r_{qc}
Mq2	30	.43	.50	.50
K2	48	.24	.16	.29
C	48	.57	.50	.58
D2	37	.07	.24	.24

TABLE III.

Group	C4 (N = 36)			D4 (N = 26)		
Coefficients	r_{ac}	r_{bc}	r_{qc}	r_{ac}	r_{bc}	r_{qc}
1st Year	.31	.49	.40	.23	.51	.51
2nd Year	.55	.57	.68	.07	.26	.23
4th Year	.24	.35	.35	.26	.62	.62

Similar results were obtained when comparisons were made with second and fourth year term examination marks. (See Tables II and III above). In each case the combined intelligence test results and Arithmetic-English examination giving higher correlations with term examination marks than either of these taken separately. The conclusions drawn from the experiment are that the intelligence test gives the best single forecast/

forecast of secondary school progress; that a more accurate prognosis is obtained when the combined test and examination results are used; that "abilities measured by the intelligence tests are not the same as those measured by the ordinary examination and that the common ground between them is not so great as might be supposed." As both the abilities measured by the tests and by the examinations are important factors in the determination of success in the secondary school it is important that a measure of each be a component of the examination to select children for secondary education.

Valentine in his investigations[1] deals mainly with

---

c.p. Reliability of Examinations.

---

the results of the entrance examinations and their prognostic value for secondary education. Four different criteria of success in the secondary school are taken. (1) The gaining of a School Certificate four or five years later, (2) the school's order of merit as determined by internal examinations, (3) the order of merit in the School Certificate Examination, and (4) the order of merit obtained by combining the school's order of merit with that given by the School Certificate Examination. The pupils who reached the School Certificate stage in the secondary school were arranged in order of merit according to their mark at the extrance examination and this list/

list was divided into three equal groups, "the top group", "the middle group" and the "bottom group".

Considering the performance of the School Certificate Examination as a measure of the pupils' success, Valentine admits, of course, that the gaining or losing of a School Certificate is no reliable estimate of a pupil's ability since the award is often separted from the failure by a mere mark or two. The results (see Table I, Appendix I) show how, in all cases, a large percentage of the candidates who were placed in the lowest group at entrance do better than many who were in the top group. There is thus the possibility that some who just fail to secure entrance to the secondary school might have been successful in gaining the Certificate.

A similar result was obtained when the matriculation examination was considered. The percentage passing the matriculation standard in the various centres is given in Table II, Appendix I.

Again, it is admitted that the criterion is not extremely reliable, the more so as the numbers at this stage are small. Nevertheless, the results do show that many pupils in the middle and bottom third at entrance ultimately do better than a large percentage of those in the top third. The pupils at the matriculation stage are of a comparatively high ability level hence it is concluded that the entrance examination fails also/

also as a method of selecting the pick of the pupils.

As has been pointed out, the gaining or losing of a School Certificate or "matriculation" may depend on a few marks in a compulsory subject, hence to take these as measures of a pupil's ability is unsatisfactory. Valentine states that "a pupil A, with total of marks 50% greater than that of pupil B may fail to gain a Certificate, though B gets one."

To avoid this uncertainty Valentine then takes as his criterion of success (1) the order of merit in the School Certificate Examination, (2) the order of merit in the schools own examinations at the time of the School Certificate Examination, and (3) the order of merit obtained by blending these two orders. The correlations between each of these three orders and the order of merit given by the entrance examination were calculated (See Table III, Appendix I). In each case the correlations were of such a magnitude as to support Valentine's conclusion that the entrance examination gives little or no prediction of the order of merit of the pupils in these examinations four or five years later.

There is a strong possibility that the chance errors in the school examinations and the Certificate Examinations will cancel each other. Added to this is the fact that there was a high degree of relationship between the order of merit in the two examinations. For four centres, the correlations were/

were ·75, ·69, ·87, and ·87, giving an average of about ·8. Valentine's conclusion is that the blended orders may therefore be taken as representing with a high degree of reliability the examination capacity of the pupils in the subjects taken in these examinations. The entrance examinations, therefore, failed to predict this examination capacity four or five years later.

Valentine's investigations deal only slightly with the use of intelligence tests as part of an entrance examination. In Centre J, which is considered as having the most reliable entrance examinations, mental tests were introduced along with other changes such as a more detailed standardisation of the marking of essays and the like. Prior to the introduction of the tests, the average correlation coefficient between entrance and School Certificate orders for fifteen classes over a period of three years was ·35, while for the two years, 1926-27, when the new conditions were in operation, the average correlation coefficient was ·46.

It is worthy of note, at this point, that a coefficient of the order of ·5 indicates marked discrepancies between the two sets of variables compared. Otis has shown[1] that with

---

1. Statistical Method in Educational Measurements. A.S. Otis. Chicago, 1926. p. 223.

---

such/

such a coefficient one third of the successful candidates should have been replaced by pupils who failed to gain admission. Indeed, the coefficients between the entrance examination and the subsequent careers of the pupils are such that it may be concluded that several who gained admission to the secondary school should have been excluded while, conversely, several of the failures would have justified their admission. Valentine concludes "that at present, there is no well-established agreement as to how best to select pupils for the type of education now usual to secondary schools."

In Centre G, the Headmaster tested the pupils on entrance with Spearman's intelligence tests and found that the tests were decidedly superior to the entrance examination. The actual figures are given in Table IV.

TABLE IV.

Correlation between	1928	1929
Entrance Examination and School %	·29	·03
Test and School %	·43	·57
Number	59	14

The 1929 results, of course, mean little due to the number in the group although the tendency in the superiority of the test is indicated. From these and reference to other investigations, Valentine concludes that the intelligence test is/

is more reliable and more valid than the entrance examination but that, on the whole, it seems safe to say already that the combination of written examinations plus carefully selected intelligence tests is likely to supply more reliable means of selecting pupils for secondary schools at 11+ than will the examination alone.

Valentine has also tried to find out the causes of the discrepancies between the entrance examination and the subsequent attainments of the candidates in the secondary school. Sources of error may be classified into three groups (1) the unreliability of the criterion (2) the unreliability of the entrance examination and (3) causes which influence the child's performance at the examination.

Considering the reliability of the School Certificate Examination and the school marks, it seems evident that the results of these examinations, including as they do the results of several examinations, would offer less opportunity for chance errors than a single examination such as normally given for entrance to secondary schools. It has been admitted, of course, that the passing or failing in the School Certificate Examination is not of itself a reliable criterion. There is the possibility that the orders of merit given by the schools and the School Certificate Examination may be unreliable. In four centres, the average correlation between these orders was ·8,

while/

while for four classes in another centre it was ·82, which shows that the two orders were to a large extent similar. When the entrance examination order was compared with the School Certificate order for two groups the coefficients were ·36 and ·37. Now, as has been indicated, the combination of the school's own order and that of the School Certificate should eliminate chance errors but it was found that this blended order gave no greater correlation with the entrance examination. The explanation of the discrepancies, therefore, must be with the entrance examination or other causes.

Another reason given for the discrepancies may be that qualities of character, industry, persistence and the like, influence the secondary school work to a greater extent. To test this argument it was decided to investigate the change within the school's own order over a period of years. Taking a group of 12 to 13 year olds and comparing their marks after a year in the secondary school with those after four or five years the correlation was ·79 (74). Dealing with a group of about 170, the following results were obtained to give a comparison between the school order from year to year.

TABLE V.

Correlation of Orders from Year to Year in the Secondary School.

	Entrance Order with 1st Year Order	1 v 2	2 v 3	3 v 4	4 (or 5) v School Certificate
With age allowance	.43	-	-	-	-
Without age allowance	.52	.74	.81	.79	.82

Table V. indicates how little variation occurs in the school order from year to year.

The theory that these high correlations are due to the fact that the teachers' marking is influenced by their knowledge of the pupils' work which tends to make the marks constant is also refuted. In the school from which the results in Table V were taken, each question of the examination papers was marked by a different teacher thus eliminating the "halo" effect. Again, the correlation between the first year mark and the results of the School Certificate Examination which were corrected by external examiners was ·63, being much the same as between the first and four year school orders, ·62 (115). The conclusion drawn from these results is "that the mere instability of the schools' own examinations and the temporary variability of a boy from year to year are the main causes of changes within the school and not a steady and continuous pulling ahead due to character qualities, or to continued and progressive growth at different rates, in spite of the fact that undoubtedly the younger able boys tend to move ahead of their older less able ones."

In one centre, it was found that when the order of merit after one year at the secondary school was compared with the order at the School Certificate stage the average correlation was ·69 (298).

The/

The corresponding result for the entrance examination and the School Certificate Examination was ·37 (302). From this it is inferred that there is a considerable change in order during the first year. To meet the criticism that a three year period is being compared with a four year period, the correlation between entrance order and school order three years later is compared with the correlation between the school order after one year and the School Certificate order. The corresponding coefficients are ·35 and ·67. This supports the argument that the big changes come in the first year in the secondary school. This is confirmed by correlating the entrance examination and the first year order, the result being ·43. This greater change of order which occurs in the first year is due probably to the new type of work done in the secondary schools including as it does Science, Mathematics, a foreign language and so on. This is a very important point as it seems to indicate that the best time for predicting success in the School Certificate Examination is at the end of the first year in the secondary school.

The discrepancies in the order at entrance and the order four or five years later may be due, also, to the unreliability of the entrance examination; the unreliability in marking and the unreliability of the examination itself. That some of the entrance examinations have a high degree of reliability is shown by the results of an experiment carried out by Valentine. The/

The English papers for 1930 and 1931 were set to a class of 40 within a week of each other and marked by two examiners A and B. The reliability of the examination when A marked the questions was ·90 and ·93 when B was examiner. A similar experiment in Arithmetic gave the reliability as ·90. The reliability of the marking can also be improved by the introduction of detailed instructions for the examiners and a board of examiners.

Other causes for the discrepancies arise when special coaching has been given to pupils sitting an entrance examination. The variation in teaching from school to school also creates differences between the entrance order and that of a few years later. Another very important point is the fact that a test in English and Arithmetic is not sufficient to predict the ability to undertake successfully the study of a wide range of subjects in the secondary school.

Another valuable follow-up investigation was that made by Collier in Northumberland[1]. For several years the

---

1. The Predictive Value of Intelligence Tests for Secondary Education. J.W. Collier. Brit. J. Educ. Psych., Feb., 1933. p. 65.

---

method adopted for the selection of children for secondary schools in this county was to submit all candidates to a battery of three tests Intelligence, English, and Arithmetic. An investigation/

investigation was made to ascertain how this admission system worked. After entrance, the progress of each pupil was recorded and an estimate of his suitability for the secondary school course was made by the head teachers on an A, B, C, D, E, rating scale. Each category had a further scaling under three classes, for example, A1, A2, and A3. To lessen the influence of the difference in standards of assessment the schools were divided in three groups. Group 1 with 75 pupils from six schools; Group 2 with 81 pupils from five schools, Group 3 with 70 pupils from three schools and Group 4 with 117 pupils from seven schools. Each pupil had completed at least three years, most of them five years, and all were over 11 and under 12 years of age at entry. In the entrance examination the scores of the three tests were adjusted so that in the composite mark each had equal weight.

The correlations between the various entrance tests and the secondary school assessments are given in Table VI.

When the partial correlation for these data were worked out the coefficients in Table VII were obtained.

The correlation between the English Test score and the secondary school assessment does not give the true relationship between these two values. This correlation may be due to some extent to the correlations between the assessment and the Arithmetic and Intelligence Tests. To remove the influence of these/

these factors the correlation may be based on a group in which all the pupils have the same Arithmetic and Intelligence scores. This can be done by the technique of "partialling out" these factors.

TABLE VI.

Correlation between Secondary School Assessment and	Values of r				Av.
	Group 1	Group 2	Group 3	Group 4	
A = Arithmetic	.33	.20	.24	.25	.26
E = English	.30	.36	.18	.42	.32
E + A	.41	.37	.27	.44	.37
I = Intelligence	.47	.47	.51	.48	.48
T = Total	.50	.51	.42	.53	.49

TABLE VII.

Correlation between Secondary School Assessment and	Partial Correlations (other two eliminated.)				Av.
	Group 1	Group 2	Group 3	Group 4	
A	.18	-.10	-.05	.01	.00
E	.15	.30	.09	.37	.23
I	.32	.42	.45	.38	.39

Table VII shows that the 'raw' correlations, with the exception of those involving the intelligence test, are considerably reduced. It may be concluded, therefore, that intelligence/

intelligence plays a large part in the ability to answer the English or Arithmetic tests or that the English and Arithmetic tests are measuring much the same ability as the intelligence test. On the other hand, the intelligence factor is fairly constant and independent of the other two.

A similar investigation was made, under less rigid conditions, on the school career of 346 pupils. The results are given in Table VIII.

TABLE VIII.

Correlation between School Assessment and	Observed Correlation Coefficients	Partial Co-efficients
Arithmetic	.34	.12
English	.37	.24
Intelligence	.45	.35
Combined Total	.51	-

As far as prediction is concerned, correlations of the magnitude given in these tables would mean but little difference from a chance selection. Collier points out, however, that he is dealing with highly selected groups which thus reduces any correlation (See "The Essentials of Mental Measurement". Brown & Thomson. Chap. VII). The effect of this selective process is seen when the correlations between the various tests are considered. These are given in Table IX (Page 49).

TABLE IX/

avoid the predominance of the Arithmetic scores due to the wider scatter of the marks in this test.

Hughes, in an investigation, sets out with the aim of discovering the reasons for the discrepancies between the Intelligence Tests and entrance examinations to secondary schools.[1]

---

1. Discrepancies between the Results of Intelligence Tests and Entrance Examinations to Secondary Schools. A. G. Hughes. Brit. J. Psych. Vol. IV, Part 3, 1934. p. 221.

The data for this investigation was obtained from the London Junior County Scholarship Examinations which were held twice yearly for children between the ages of 10 years 6 months and 11 years. The examination consisted of an English Essay, papers in English comprehension, mechanical Arithmetic and Arithmetic problems. Scholarships were awarded on the results of the examination and an intelligence test was used to supplement the examination results. The correlation between the examination and the test results was fairly high in comparison with those of previous investigations. Data for two separate years are given in the following Table.

TABLE XI.

	1927		1929	
	r	N	r	N
Boys	.64	436	.66	1107
Girls	.59	394	.62	1083

A correlation of ·63, which is a rough average of the correlation in the above table, allows of a fairly large number of discrepancies between the examination and the test. Table XII gives the awards by means of the examination alone and the number (and %) of these who would have gained scholarships by means of the test alone.

TABLE XII.

	1927		1929	
	Boys	Girls	Boys	Girls
On Examination alone	43	39	106	95
On Test alone	21 (49%)	23 (59%)	55 (52%)	51 (53%)

Dr. Ballard made an enquiry into the discrepancies between the 1927 examination and test. He found that "in some instances the scholarship examination fails to capture quite brilliant children." With regard to the 1927 examination, 120 striking cases of discrepancy were selected, 30 boys and 30 girls who won scholarships by means of the examination and the same number who would have won scholarships by means of the intelligence test. Of these, 40 cases were investigated personally by the author while the remaining 80 were 'followed-up' by means of a questionnaire sent to the schools concerned.

That the entrance examination is not an entirely satisfactory method of selection is shown by the secondary school teachers' estimate of the ability of the scholarship pupils/

pupils. The following table gives a percentage number of children judged fit, doubtful and unfit.

TABLE XIII.

	Boys	Girls	Boys & Girls
Of Scholarship Standard	42	35	38
Possibly of Scholarship Standard	21	42	32
Definitely not of Scholarship Standard	37	23	30

Hughes gave three conditions as necessary for success in the secondary school. (1) A good degree of general intelligence (2) a certain standard of academic attainment on entry and (3) some sturdiness of character preferably combined with physical vigour and satisfactory home conditions. This seems to indicate, then, the use of an intelligence test and an entrance examination along with a rating by the teacher and the schools' medical officer.

The intelligence test is required as some children may fail through extraneous causes to reach the necessary level of attainment yet possess the necessary intelligence which may be revealed by the test.

Extraneous handicaps which militate against the intelligent child gaining a scholarship are bad home conditions; defects of character such as laziness, carelessness and lack of concentration; poor health; irregular attendance; ineffective teaching/

teaching and the unreliability of the examination. Hughes gives a frequency distribution of extraneous handicaps under the headings home, health, character and attendance, both for the pupils who gain a scholarship on the results of an intelligence test alone and those whose awards are based on an examination alone. (See Table IV, Appendix I). The table illustrates clearly a large number of intelligent children who are prevented by such handicaps from gaining scholarships which are planted by means of en entrance examination alone.

An analysis of the home conditions is also given for both groups of scholarship winners. (See Table V, Appendix I). Here it seems to be indicated that a child must come from a reasonably good home before he has much chance of gaining a scholarship by means of an entrance examination.

A second reason for the inclusion of an intelligence test is that some children may achieve the necessary level of attainment yet not possess sufficient intelligence to undertake successfully a secondary school course. Reasons for this may be that the child enjoys some special advantages such as good home conditions, some merit of character such as ambition, the benefit of good teaching or special coaching. Another reason may be found in the unreliability of the intelligence test. Even though an intelligence test has a reliability of the order of ·9 this means that there are considerable discrepancies/

pancies. Again, it may be that a successful scholarship candidate did not do himself justice in the test.

Defects in character shown in the elementary school by a child, no matter how intelligent, if they persist in the post-primary school, would prevent the successful undertaking of a secondary school course. In order to find out whether character defects developed by the age of 11 are permanent, Hughes found the percentage number of children reported as suffering from character defects both by the elementary and by the post-primary school teachers. (See Table VI, Appendix I). From his data he concluded that "it is reasonable to suppose that since most of these fects in the extremely bad cases are presumably the result of permanently bad conditions beyond the influence of the schools, the chance of their disappearing is small."

Another point made in this investigation is that the teachers' forecasts of scholarship success are in closer agreement with the results of a scholarship examination than with these results based on an intelligence test. (See Table VII, Appendix I). This might have been expected as the teachers' forecasts will have been based on the result of examinations similar to the scholarship examination.

The general conclusion drawn by the investigator is one which resembles closely that of several of the others, namely/

namely, "that neither an intelligence test nor the customary examination, is by itself, a satisfactory method of selecting children for secondary schools, but that a carefully constructed intelligence test should be added to the usual type of examination in English and Arithmetic."

Sandon has made a study of the unreliability and invalidity of examinations.[1] He suggested that there were

---

1. The Necessary Imperfections of an Examination. F. Sandon. Brit. J. Educ. Psych. Vol. V, Part II, June 1935.

The Predictive Value of School Examinations and Psychological Tests. F. Sandon. Transactions, British Association, 1933. Sections J & I.

---

twelve steps in the normal examination procedure and at every stage there is a liability of error, biassed or random. His final conclusion is that "If all fluctuations coexist, then we shall be fortunate if our examination correlates ·6 with the criterion: it will be a remarkably good examination where the correlation is ·8 with the criterion, and it will be one altogether exceptional and beyond the reach of ordinary practice if the correlation were ·95."

Professor Thomson has also carried out a follow-up experiment with the results of Moray House intelligence and scholastic tests given in the West Riding as the examination for free places in the secondary school.[2] The results are obtained from the

---

2. The Value of Intelligence Tests in an Examination for Selecting Pupils for Secondary Education. G.H. Thomson. Brit. J. Educ. Psych., Vol. VI, Part II, June 1936.

---

marks of 613 candidates distributed among ten secondary schools. Headmasters submitted orders of merit for the pupils two years after entrance. The correlations between these orders and the various entrance tests are given in the following table:-

TABLE XIV.

Headmasters' Order of Merit and	r
I.Q.	.410
A = Arithmetic	.379
E = English	.346
A + E (equal weight)	.445
A + E + I.Q. (equal weights)	.491
Best weighted team of A, E and I.Q.	.492

The best weighting of the battery of tests was I.Q. : A : E = 1 : ·91 : ·67. Thomson concludes from these results that "the three parts of the examination should be given equal weight, with increasing weight to the I.Q. as the border line is approached (to avoid ties), since it is slightly the best single measure of the three."

Thomson points out, also, just as Collier has done, that he is dealing with highly selected groups and selection decreases the correlation. The scatter of the intelligence in the secondary school pupils being only three-fifths of that of the/

the whole child population and that, if the group had been made by a purely random selection, then the correlation would have been of the order of ·8. It is worth observing, also, that the correlations were calculated separately for each school, and their weighted average found.

The inter-correlations between the three parts of the entrance examination are shown in Table XV.

TABLE XV.

r_{AE}	.326
r_{AI}	.409
r_{EI}	.438

These inter-correlations differ considerably in magnitude from those given by Collier and in the fact that the Arithmetic test correlates to a less extent with the intelligence test than with the English test.

Earle[1] has approached the problem in a different

1. Tests of Ability for Secondary School Courses. F.M. Earle. University of London Press, 1936.

manner from most of the other investigators claiming that what is necessary is not so much a test of general ability but tests which will predict the child's success in the various subjects of the secondary school course. A reliable test of general intelligence/

intelligence is useful for forecasting all round proficiency especially with reference to scholastic examinations of the Leaving Certificate type. But less than 10% of the school population reaches this stage and thus what is of greater importance is what he terms "differential diagnosis". By this he means the measuring of the "capacity of an individual to reach a high stage of proficiency in a single subject or a group of allied subjects."

With a group of pupils, all of whom achieved a minimum "pass mark" of 55 in an entrance examination, the correlation between the English part of the examination and the pupils marks in Language subjects (English, Modern Language, Latin) five years later was .64. That between the Arithmetic part and the fifth year Mathematics marks was .41. Actually the English section of the entrance examination has a better predictive value as regards fifth year Mathematics and Science marks than the Arithmetic section as the correlation for the former is .57.

When the total marks in the entrance examination were correlated with the total marks in the fifth year, the coefficient was found to be .64. The children of this group were tested some time later with an intelligence test which was found to correlate to the extent of .48 with the entrance examination and .49 with the fourth and fifth year total class marks.

It/

It would appear, then, that here is further evidence in support of the idea that, at present, the most successful method of selection of pupils for the secondary schools is a combination of an entrance examination and an intelligence test. Earle, of course, puts forward the case for the discovery of special rather than general ability required in the success which is to be predicted.

A rather interesting point is made in connection with prediction of success for which a correlation coefficient must be high before the prediction can be made with any degree of accuracy. On the other hand, Earle claims that it may be possible that a more accurate prediction of failure may be made from the same data. The method suggested would be to find a mark to indicate the minimum standard of capacity required for ultimate success. Cases are quoted of candidates who failed to achieve such a mark in one entrance examination and whose secondary school record is unsatisfactory. Of eleven, in this group, who were retarded at some stage of their secondary course, nine obtained marks below the minimum standard.

It must be admitted that tests of special abilities such as Earle suggests are desirable yet the intelligence test gives a forecast of all round learning capacity. Although success at the Leaving Certificate stage depends on temperament and environment as well as innate capacity, yet a measure of this/

this capacity has a value. In the following table Earle shows, for example, the relationship between the intelligence test results and the success in the Leaving Certificate Examination.

TABLE XVI.

Mental Ratios of Leaving Certificate Candidates

Mental Ratio	Chance of Success at Leaving Certificate Examination	
	Fifth Year	Sixth Year
120 and over	62%	96%
112 to 119	30%	83%
Less than 112	9%	17% (7th Year 50%)

A similar enquiry gave the following data.

TABLE XVII.

	No of Cases	Mean Mental Ratio
Awarded a Group Leaving Certificate in Fifth Year	19	116
Awarded a Group Certificate in Sixth Year	4	106
Failed on Presentation in Fifth Year and left school	6	106
Not presented in Fifth Year and left school	6	98
Left before reaching stage of presentation	5	112

The/

The mental ratios in the previous table are lower than in Table XVI but the relative differences are similar. The following prediction on the basis of the intelligence test data are given.

5 were successful in 5th year. (agreeing with the forecast).
33 were unsuccessful in 5th year. (agreeing with the forecast).
3 out of 6 who were said to have a 50% chance of success in the fifth year were successful.

Scanty as the data is, it does prove that, though perhaps not by itself but along with other information such as school records, the intelligence test can play a part in the selection of children for secondary education in which the scholastic examination plays such a large part in the determination of success.

In connection with the selection of children for secondary education mention may be made of a recent book which reviews the methods used by the various committees in England, "The Selection of Children for Secondary Education."[1] The authors, who have had wide experience in

1. The Selection of Children for Secondary Education. J. B. T. Davies and G. A. Jones. Harrap. 1936.

examining for county scholarship examinations, give as the list of the types of tests used (1) Written examinations in English and Arithmetic (2) oral examinations (3) Intelligence and Scholastic tests (4) internal school tests (5) primary school records/

records, including the estimates of the Heads of the schools.

Although no definite experiment has been carried out, the authors discuss each test in turn and give a description of the modern techniques which have been devised to increase the reliability of prediction. Standardized marking, the construction of the examination papers, age allowance and so on are dealt with in turn.

Summing up the whole situation they conclude that the basis of all secondary school entrance examinations (for England) is a test, of some type or another, in English and Arithmetic. Also, from a study of research, "the prognostic value of the intelligence test is higher than that of the ordinary written test in either Arithmetic or English, but is sometimes very little higher." And "the highest prognostic value is obtained from carefully devised and mathematically controlled combination of intelligence test and written examination in Arithmetic and English." To ensure any measure of reliability in a selective examination, they claim, that the results of an intelligence test are essential.

With regards teachers' estimates, and school records, these are considered to be unreliable though they may be valuable for differentiating between pupils from the same school. One headmaster is quoted as having estimates which never correlate less than .8 with the School Certificate Examination marks.

In/

In May, 1933, the Edinburgh Education Committee approached the Education Department of Edinburgh University with a view to finding out the comparative merits of a Moray House battery of tests, a Qualifying Examination in English and Arithmetic, and Teachers' Estimates. The testing took place in January and June, the tests being the same as those used in th Dundee enquiry, namely, M.H.T. 20, M.H.E. 7, M.H.A. 7, M.H.T. 21, M.H.E. 8, and M.H.A. 8. The report has not been published but was issued privately by the Education Committee.[1.]

---

1. Edinburgh Corporation. Education Committee. Report on Experiment in the Use of Standardised Tests, Conducted by the Education Department, University of Edinburgh, October, 1934.

One of the first points which arises from a study of the distribution of the marks is that in the teachers' estimates and the qualifying examination there is insufficient "headroom". In other words, there is little or no differentiation between the abler pupils. The tests, however, were constructed so as to give both sufficient "headroom" and "footroom".

A second point is also made, namely, that the intended weighting of two parts of an examination is not always operative. For example, in the Qualifying examination the possible marks for English and Arithmetic were 180 and 120 respectively. The standard deviations, however, upon which the weighting of examination marks depend were 26·8 and 25·3 in December and 26·1 and/

and 24·4 in June. This means that in the total Qualifying examination mark these subjects are really given almost identical weight although the apparent plan was to give the English mark 1½ times as much weight as the Arithmetic.

Unfortunately, the report was issued before any follow-up data was available. In order to find out what test has the greatest value, an average for all the tests was found for each pupil, it being decided that this average mark of seven subjects would be a fairly good estimate of the pupil's ability. The correlations of the tests with this average mark are given in Table XVIII.

TABLE XVIII.

Correlation Between Average Mark and	r
English Records	.864
Intelligence Test	.859
Arithmetic Test	.851
Arithmetic Record	.837
Arithmetic Examination	.835
English Examination	.830
English Test	.811

Actually, there is little to choose between the tests as all the coefficients are very close to one another. As it is/

is more usual in actual practice to take these marks in combination the English with the Arithmetic, correlations were worked out with the following result:-

Scholastic Tests (E + A)	.926
Qualifying Examination	.926
Teachers' Estimates	.915

All three tests together, Intelligence plus English plus Arithmetic, gave a correlation of ·945. Although the final conclusion is that the tests are a little superior to the Qualifying examination and the teachers' estimates, yet it must be emphasised that the only valid method of determining this fact is by following up the subsequent careers of the children. To find out how the various combinations of tests agree as far as arranging the pupils in order are concerned the following multiple correlations were found.

Teachers' Opinion (Te + Ta) against Qual. Exam.(Qe + Qa)	.812
Teachers' Opinion (Te + Ta) against Stand. Tests(E + A)	.578
Qual. Exam. (Qe + Qa) against Stand. Tests. (E + A)	.796
Teachers' Opinion (Te + Ta) against Intelligence Test	.703
Qual. Exam. (Qe + Qa) against Intelligence Test	.715
Stand. Tests (E + A) against Intelligence Test	.816
Stand Tests (I + A + E) against Teachers' Opinion + Qual. Exam.	.826

Teachers' estimates generally suffer from the fact that the standard of marking varies from school to school and that when the results of several schools are slumped there is a lowering of the correlation. This point is illustrated by the following table:- (See page 66)

TABLE XIX/

TABLE XIX.

	1st school	2nd school	3rd school	4th school	All four schools
English Records with English Exam.	.93	.89	.90	.95	.835

This difference of standard in different schools is one of the biggest objections to complete reliance for promotion purposes on teachers' opinions alone.

Summarising the results of these British investigations it may be said that the early reports and experiments dealt mainly with the introduction of the intelligence test as a means of selecting pupils for secondary education. It was found that such tests might reveal ability or the lack of ability not otherwise evident - in other words, the intelligence test was a valuable supplement to the examination. Some authorities granted several awards solely on the results of an intelligence test and, in most cases, the subsequent secondary school careers justified these awards.

At first, there were no actual experiments. Investigators based their conclusions on specific cases. Later investigations conducted on scientific lines justified these early conclusions; Amos, Valentine, Collier, Hughes and Thomson, to mention but a few, concluded that the intelligence test/

test was the best single measure for predicting post-primary school success.

Even from the first, it was admitted that an examination should be retained. The reasons for this were twofold, (1) a certain level of attainment is necessary before a pupil can profit by a secondary school course, no matter the level of intelligence (2) an examination keeps a uniform standard over an area.

The early investigators realised, also, the importance of the borderline in a selective examination. On the whole, it is easy to select pupils who are sure to do well in a secondary school and those who would obviously make little success in the course. The great difficulty lies in the selection from the intermediate group. Oates and Valentine, for example, showed how unreliable an examination could be at this point although later investigators have not developed this aspect as fully as might have been done.

Several investigations were made into the causes of the discrepancies in the customary examination in English and Arithmetic. The different standards of teaching in schools, home conditions, special coaching, and the unreliability of the examinations themselves, are among the most important causes. Perhaps the most important of all is that an examination in English and Arithmetic is not sufficient to predict the special abilities/

abilities which are required for the successful undertaking of a secondary school course.

Teachers' estimates were found to give an unsatisfactory prediction of success. The chief reason for this seems to be that the standards vary so much from school to school. Another reason is, of course, that these estimates are generally based on school examinations. To overcome the first difficulty some scaling of marks might have been attempted although there seems to be no published record of such an experiment.

The majority of the investigators found that the intelligence test gave a better prediction of success in the secondary school than an examination in English and Arithmetic and a combination of the two was better than either. Collier and Thomson, however, pointed out that a weighted combination of intelligence, English and Arithmetic tests gave the best prediction of all, increasing weight to be given to the intelligence test as the borderline is approached.

Valentine and Earle, showed that the results of the first year in the secondary school gave a fairly reliable prediction of a pupil's success four or five years later. This implies a probationary period in the secondary school although the length of such a period seems to be a matter of doubt. Table XX gives a summary of the correlation coefficients found in the chief British investigations.

TABLE XX/

TABLE XX.

Correlations Between Secondary School Marks and other Factors.

.	I.Q.	Qe + Qa	E + A	I + E + A or I + Q	Weighted Combination of I, E & A
Amos	.40	.33		.45	
Valentine	.43	.30			
Collier	.48		.37	.49	
Thomson	.41		.45	.491	.492
Earle		.64			

C. AMERICAN INVESTIGATIONS.

In America, research on the question of the progmostic value of tests at the entrance to the junior high school is also scanty. One explanation of this is that most of this work has been done at the University stage mainly because the University authorities were primarily interested in the problem as dealing with themselves and were the only bodies who could supply the necessary money for such research.

It is the elementary school which decides whether a pupil will be fit for promotion to a high school. Sometimes this is done by means of a special examination and sometimes by the school record but the high school takes little interest in the problem of admission. The result of this, it is held,

is/

is that 20% of those entering drop out before the end of the second year and thus the problem is one of guidance rather than admission. Some means must be available to predict school success although, as has already been pointed out, the definition of success in school raises many problems. On the whole, however, the generally accepted criterion is school marks.

One of the earliest investigations was that by W.R. Miles in 1911[1] in which he compared the average

1. Comparison of Elementary and High School Grades. W.R. Miles. Iowa City, University of Iowa, 1911.

elementary school marks and the average of high school marks of those who had been in the high school for two to four years. The correlation between the two was ·71.

The next outstanding research was that by T. L. Kelley[2]

2. Educational Guidance. T. L. Kelley. New York, Teachers' College, Columbia University, 1914.

he criticised Miles's results from two standpoints (1) the correlations would have been larger if only first year results in the high school had been taken, for, in general, as the time between the testing is increased the correlation decreases (2) the greater the number of cases the more reliable is the average. It is doubtful as to which of these factors is of greater weight. In his own investigation, Kelley found that the/

the best prediction is obtained from elementary school marks, next, teachers' ratings, and least of all from specially devised tests. (See Table VIII, Appendix I).

A result differing from that of Kelley's was obtained by Fretwell[1] who found that the highest correlation was

---

1. A Study of Educational Prognosis. E. K. Fretwell. Teachers' College Contributions to Education, No. 90. Teachers' College, Columbia University, New York, 1919.

---

obtained from a composite of eleven tests. (See Table IX, Appendix I). It is worth noting, however, that Fretwell's tests gave only a slightly better prediction than Kelley's tests (·56 against ·51) whereas in Kelley's investigation the elementary school marks had a much better predictive value than in the one conducted by Fretwell (·79 against ·49).

Several researches were made to investigate the value of the intelligence test as a means of predicting success in the high school. In 1923, Jordan[2] reviewed the typical

---

2. The Validation of Intelligence Tests. A. M. Jordan. J. Educ. Psych., Vol. 14, 1923. pp. 348 - 66, 414-28.

---

studies on the relationship between group intelligence tests and high school marks. He found that the average of 50 different coefficients was .43. His own results on 67 first year/

year high school pupils show correlations averaging ·47.[1]

1. Correlations of Four Intelligence Tests with Grades. A.M. Jordan. J. Educ. Psych., Vol. 13, (1922) pp. 419-29.

Proctor[2], in his investigation obtained correlations

2. Psychological Tests and Guidance of High School Pupils. W. M. Proctor. J. of Educ. Research. Monographs No. 1, Public School Publishing Company, Bloomington, Illinois. 1923.

between first year high school marks and intelligence tests ranging from ·343 to ·545.

Ross[3], Brooks[4], and Fleming[5] also made investigations

3. The Relation Between Grade Schools Record and High School Achievement. C. C. Ross. Teachers' College Contributions to Education No. 166. New York. 1925.

4. Sectioning Junior High School Pupils by Tests and School Marks F. D. Brooks. J. of Educ. Research, Vol. 12 (1925), pp. 359-69.

5. A Detailed Analysis of Achievement in the High School. The Comparative Significance of Certain Mental, Physical and Character Traits for success. C. W. Fleming. Teachers' College Contributions to Education, No. 196, Teachers' College, Columbia University, New York, 1926.

into the value of the I.Q. for the prediction of success in the high school.

Table XXI (See page 73) summarises results of all these investigations. The correlations range from ·34 to ·60 with an average of ·47. This means of course that the I.Q. has little predictive value for high school success, although it/

it must be remembered that these coefficients are based on small numbers. It is rather surprising, also, to find that the Binet I.Q. is little better than the I.Q. from Group tests as the former is much more reliable and gives a truer indication of the pupils innate ability.

TABLE XXI.

Correlations Between High School Marks and Intelligence Test Results.

.	Jordan's Summary	Jordan	Proctor	Ross	Brooks	Flemming
Binet I.Q.			.545 .487		.474	
Army Alpha	.38	.476	.413			
Army a			.343			
Miller	.56	.476			.501	.47
Otis	.47	.450			.541	
Terman	.47	.492		.37	.401	.60

The relative value of standardised tests and elementary school records in predicting high school success was also the subject of several investigations. Ross, for example, supports Kelley in claiming that the elementary school record has a greater predictive value than achievement tests. The correlations with the marks in the first year in the high school are as follows:-

Grade School Composite	.63
Terman Group Test of Mental Ability ...	.37
Thorndike - McCall Reading Test	.33
Woody - McCall Arithmetic Test	.40

In/

In his investigation, Ross found the predictive value of the separate subjects in the elementary curriculum (See Table X, Appendix I) and thus was able to tell which were the most important factors in predicting high school success. By taking a weighted average of selected subjects he obtained correlations of ·68, ·67, ·56 and ·65 for four successive years which was only a slight improvement on the result with the simple average of all subjects, ·60.

The correlation between composite scores obtained from grade school factors, and average standing of pupils who remain in high school for two or more years were found to be ·60, ·64, ·57 and ·66 respectively for four successive years.

Brooks, in his book "The Psychology of Adolescence"[1], claims that "the best single basis for predicting scholarship (i. e. school marks or other measures of scholastic attainments) in high school is the average mark received in the grades immediately preceding the high school. Next in order of predictive value are teachers' estimates, (of industry, school attitude, intelligence, persistence, conscientiousness, etc.) intelligence and achievement tests and chronological age."

1. The Psychology of Adolescence. F. D. Brooks. Harrap & Co. Ltd. pp. 553 - 574.

In his investigation, Brooks considered two criteria as measuring success in the high school. First, the average of/

of the marks in Mathematics, History, Geography and English for two 'semesters' and, second, a composite of the former and the scores on the Stanford Achievement Tests.

The results of this investigation also support the argument that the elementary school record is superior to tests in predicting success in the junior high school. For a group of 93 pupils the correlation between the marks in the high school and the elementary school is ·70 while the corresponding figure for the tests is ·50.

Brooks also realised the possibility of combining different measures. The following multiple correlations were obtained:-

Elementary School Marks with I.Q's.....	.79
Elementary School Marks with Achievement Tests.....	.76
Elementary School Marks with I.Q.s and Achievement Tests	.85
Elementary School Marks with Age	.76

His conclusion is that a composite of sixth grade marks and the results of a good group intelligence test give better prediction of average marks in the first year of high school than do any other two factors.

By this method of combining measures it is possible to obtain a higher degree of accuracy in prediction. Brooks found that he could obtain multiple correlation coefficients of ·80 or ·85 between marks in the first year of the junior high school and a composite of elementary school marks, intelligence test scores, teachers' estimates and chronological age. He suggested that/

that, if more factors were used, it might be possible to push the correlation up to ·90 although he admitted that highly reliable measures would be necessary to secure such coefficients. In this research, also, there is raised the question of the value of the average of the marks in the first year of high school. These, according to Brooks, give a good indication of what the pupil will do in the following year. The correlations between three successive years were ·64, ·69 and ·75.

Flemming's investigation is not truly one of prognosis as all the data were obtained during the year in which the marks were given thus making the correlations higher than they would have been if the ratings had been given during the previous year. It is interesting to note, however, that the highest correlation with high school marks is given by the teachers' estimate of intelligence. (See Table XI, Appendix I). Multiple correlations of ·78 and ·85 were obtained using teachers' estimates, intelligence tests and chronological age.

Several of the investigators found the correlation between success in high school and the chronological age of the pupils. The consistency of their results leads to the conclusion that this correlation is negative to the extent of about -.35.

The following gives a summary of these

results/

results:-

Miles	-·31
Kelley	-·34
Ross	-·36
Brooks	-·49
Flemming	-·30
Average	-·36

From these American investigations it may be concluded that the elementary school record is the best individual measure for forecasting success in the high school. The results of the investigations vary a great deal particularly where the data is taken from more than one school. This is due, probably, to the different standards of marking in the various schools. The average correlation between elementary school records and marks in the first year of the high school is about ·65.

The American investigators, like the English, found that the intelligence test was superior to achievement or prognosis tests for predicting high school success but not as good as the elementary school record.

The best means for predicting high school success was a combination of intelligence test results and the elementary school record. By combining the various measures multiple correlation/

correlation coefficients of about ·90 were obtained.

First year marks in the high school give a good indication of a child's future performance in the course. This result, also, is similar to that of the English investigations. Teachers' estimates, although two of the experiments gave high results (Kelley ·72 and Fretwell ·80), generally speaking proved to be unreliable.

Age correlates with high school marks to the extent of about -.36.

A summary of the results of the main experiments are given in Table XXII.

TABLE XXII.

Correlation Between First Year High School Marks and Other Factors.

Investigator.	Elem. School Marks	Teachers' Estimates	Tests	I.Q.	Est. of I.Q.	Multiple Correlation
Miles	.71				.72	.89
Kelley	.789	.76	.51			
Fretwell	.49		.56			
Jordan				.48		
Proctor				.52(Binet) .38(Group)		
Ross	.64		.37	.37		
Brooks	.70		.50	.47(Binet) .48(Group)		.85
Flemming		.70		.51	.80	.85

The/

The lack of validity of tests and examinations in the prediction of success in post-primary education may be due in part to the fact that certain specific abilities are necessary for success in the study of subjects such as foreign languages or mathematics. There is some doubt whether such abilities are sufficiently matured when the child is in the primary school to be revealed by any specially devised tests.

There have been several attempts to devise such tests although these have not been too successful. Clem tried to devise a Latin prognostic test[1]. The correlation of this test with I.Q., age and high school marks in Latin were ·48,

1. Detailed Factors in Latin Prognosis. O. M. Clem. Teachers' College Contributions to Education No. 144. Teachers' College, Columbia University, New York, 1924.

-·38, ·84 respectively.

A similar attempt has been made by Allen whose investigation was probably the first to use the full technique for the construction of a prognostic battery.[2] He obtained

2. A Study in Latin Progmosis. W. S. Allen. Teachers' College Contributions to Education No. 135. Teachers' College, Columbia University, New York, 1923.

correlations of about ·57 with the first 'semester' Latin marks falling to ·47 with the second 'semester' mark.

Kelley in "Educational Guidance" gives correlations

between/

between prognostic test results and the corresponding subjects in school.

Algebra.....	.47
Geometry.....	.42
English.....	.44
History.....	.31

Prognostic tests of mathematical ability were devised by Rogers, the battery consisting of tests of algebraic ability, geometric ability and language ability.[1] Although the

---

1. Experimental Tests of Mathematical Ability and their Prognostic Value. A. L. Rogers. Teachers' College Contributions to Education No. 89. Teachers' College, Columbia University, New York, 1918.

---

correlations are comparatively high ranging from ·62 to ·92 there are two criticisms of the test. (1) The Tests measure different kinds of abilities, algebraic and geometric abilities being perhaps specialised. (2) The tests are not truly prognostic as some knowledge of the subject is necessary before the test can be attempted.

Flemming also correlated tests and teachers' estimates with individual subjects but found the correlations of about the same magnitude as those obtained with the average school marks. Jordan correlated intelligence tests with individual subjects while Ross worked out multiple correlation coefficients for predicting success in English, Latin and Mathematics, from his elementary school data.

The/

The general conclusions which may be drawn from the above investigations are as follows.

Factors such as intelligence and elementary school marks correlate less highly the special subjects than a general average of subjects.

To predict success in a special subject one would use data that has something in common with the particular subject being predicted. For example, to predict success in English in the high school weight would be given to the elementary school marks in English.

On the whole, the prognostic tests which have been devised are not truly prognostic as some knowledge of the subject is required before they can be attempted.

D. INVESTIGATIONS IN COUNTRIES OTHER THAN BRITAIN AND AMERICA.

Collmann & Jorgensen, Australian investigators, have published a description of two experiments carried out along similar lines to that of Valentine.[1]

---

1. The Prediction of Scholastic Success. R. D. Collmann and C. Jorgensen. Educational Research Series, No. 35. Melbourne University Press, 1935.

Collmann in his report "The Prognostic Reliability of Certain Secondary School Entrance Tests", bases his results on the application of English, Arithmetic and Intelligence tests over/

over a period of four years to about 250 entrants each year to the Melbourne Boys High Schpol. Two years later, at what is known as the intermediate stage, the average of seven or eight subjects was taken as an estimate of examination success at this stage. A similar method of estimation was adopted at the Leaving Examination. In this particular school, the two examinations were conducted by the teachers themselves and the system of marking and standards were fairly uniform.

The correlations between the orders at entrance and that given by the Intermediate Certificate examination over four successive years varied only slightly and the averages are given in the following table.

TABLE XXIII.

Type of Entrance Test	School Intermediate Examination		
	Correlation	Probable Error	Coefficient of Alienation
English - Arithmetic Test	.56	.038	.17
Intelligence Test	.60	.035	.20
Combined Test	.65	.032	.24
Number of Cases	620		

The conclusions drawn from these results are similar to those of Valentine and other English investigators, namely, that the intelligence test gives the best individual prediction of success but that the highest correlation is obtained by combining an intelligence test result with the result of an examination/

examination in English and Arithmetic. It is interesting to note that the coefficient of ·65, which is that given by the combined test, means that errors made in predicting the order of merit at the School Intermediate examination from the order of merit given by the combined test will be reduced 24% when compared with those made by a chance selection. It is shown later, however, that by the use of this combined entrance test it is possible to reduce the percentage of pupils failing to obtain the Intermediate Certificate by almost 50%. Some method of reducing the percentage of failures is badly needed as, for the four year period under consideration, 45% of the entrants failed to gain the Intermediate Certificate, a percentage which would be raised to 50 when those who leave owing to unsatisfactory reports are taken into account. Jorgensen points out that by a method of weighting the various tests the coefficient of ·65 would likely be increased.

Comparing the entrance test orders with that given by the School Leaving Examination the average correlations for four successive years are given in the table.

TABLE XXIV.

Type of Entrance Test	School Leaving Examination		
	Correlation	Probable Error	Coefficient of Alienation
English Arithmetic Test	.34	.084	.06
Intelligence Test	.53	.067	.15
Combined Tests	.56	.063	.17
Number of Cases	208		

The/

The results are similar to those at the Intermediate stage though in each case the coefficients are somewhat lower, the best correlation giving an improvement of only 17% better than a chance selection. Another point to be noted is that the English - Arithmetic test has deteriorated in its predictive value to an even greater extent than the others, having only a 6% better selective value than pure chance.

Another set of correlations given in the report are those between the examination results at the end of the first year in the High School and those at the Intermediate and Leaving stage. These are given in Table XXV.

TABLE XXV.

Year	Intermediate	Leaving
1929	.90 (153)	.83 (50)
1930	.86 (149)	.88 (52)
1931	.92 (166)	.86 (51)
1932	.84 (152)	.85 (55)

In comparison with the other results these correlations are high and indicate again that the results at the end of the first year in a secondary school give a good indication of what is likely to happen throughout the secondary course.

By taking merely whether the pupil passed or failed the Intermediate examination, an inquiry was made into whether a minimum mark in the entrance combined test could be fixed so as to forecast accurately whether the pupil would pass or fail in the/

the Intermediate Examination. Adjudging this minimum mark to be 60% the results given in Table XXVI were obtained.

TABLE XXVI.

Boys obtaining 60% or less on the C.E.T. and who failed in I.C.	98	Total No of Failures
Boys obtaining more than 60% on the C.E.T. and who failed in I.C.	103	201
Boys obtaining 60% or less on the C.E.T. and who passed	7	Total No of Passes
Boys obtaining 60% or more on the C.E.T. and who passed	412	419

The original 201 failures would have been reduced to 103 if the combined entrance list had been used as the selective process. This reduction in error of about 50% justifies the use of the combined entrance test which might be made more efficient still with further improvements in the construction of these tests.

Jorgensen in his investigation on "Intelligence Tests and Entrance Examinations as Instruments for Selecting and Grading Students" sets out to discover the differences in the predictive value of the examination and the intelligence test. The further question immediately arises whether the two should be combined and in what manner this combination should be made.

The investigation was carried out in two schools of different types, Collingwood Technical School and University High/

High School, with first and second year pupils. The Entrance examination in English and Arithmetic was that in general use in these schools while the intelligence test was one devised by the investigator himself. Although the internal school examinations differed in some respects, these were taken as a criterion against which both the intelligence test and the entrance examinations were compared. In most cases the first term examination was used although for a few cases the third term examination was substituted and, in the case of the High School where monthly tests took the place of examinations, an average of the first three monthly tests were used.

The Comparative Correlations for Intelligence tests and Entrance Examinations are given in Table XXVII.(See page 87).

From this table it is concluded that the intelligence test, on the whole, has a better predictive value than the entrance examination although the difference is slight. It is worth noticing that most of these coefficients are extremely low and have but little real predictive value.

From an examination of the mean intelligence scores and entrance examination marks of the different classes the weakest sections are found to be C7-12, L IV B, U IV C and U IV D. In each of these sections the intelligence test is definitely superior to the entrance examination. This point is further investigated by a study of the correlation scatter diagrams.

To/

To take but one example, the results for Class C 7-12, 1933, are given in Tables XXVIII and XXIX. (See

TABLE XXVII.

Date of Study	Criterion School Exams.	Grade or Section	Intelligence Test	Entrance Exam.	Difference	N
1933	Term i	C 1-6	.73	.63	.10	120
		C 7-12	.52	.33	.19	105
	Term iii	C 1-6	.57	.53	.04	92
		C 7-12	.44	.36	.08	91
1934	Term i	C 1-6	.51	.49	.02	144
		C 7-12	.29	.18	.11	126
	" "	M 1-2	.43	.46	-.03	41
		B 1-6	.55	.51	.04	86
1934	Term i	L IV A	.38	.45	-.07	42
		L IV B	.59	.59	.00	33
		L IV C	.64	.74	-.10	42
1934	Term i	U IV A	.40	.34	.06	41
		U IV B	.29	.42	-.13	41
		U IV C	.47	.50	-.03	42
		U IV D	.27	-.29	.56	35

TABLE XXVIII.

Intelligence and Term I.

Criterion	Intelligence Test Score				
	46 -51	38 -45	20 -37	22 -29	12 -21
710 - 800	1	6	1	1	-
620 - 709	2	10	6	3	-
530 - 619	2	9	19	8	3
440 - 529	-	-	8	14	1
320 - 439	-	-	1	6	2

TABLE XXIX/

TABLE XXIX.

Entrance Exam. and Term I.

Criterion	Entrance Exam. Score				
	85 -90	70 -84	55- -69	40 -54	25 -39
710 - 800	5	1	1	1	-
620 - 709	4	8	7	3	-
530 - 619	2	16	18	2	3
440 - 529	1	6	5	9	4
320 - 439	1	2	3	1	2

From these tables it can be seen that eight of the nine 'failures' are below average in intelligence while only three of the nine were below the average in the entrance examination. Several tables of this nature and correlation scatter diagrams lead to the conclusion that "the intelligence test is the better instrument of selection as far as those of less ability are concerned and it is for this group that effective means of selection are most needed "since the main purpose of an entrance examination is that it shall exclude the pupils who are least likely to be successful in school. Also "in one respect the entrance examination does seem to have a definite superiority: it appears to pick out the outstanding pupils better than does the intelligence test."

In/

In combining the results of the intelligence test and the entrance examinations a weighted sum of these gives a higher correlation with the criteria than does either of these separately. The best weighting is not always convenient and rough approximations to these give fairly satisfactory results as is shown in Table XXX.

TABLE XXX.

Effect of Simple Methods of Combining Scores

Section	Correlations with Criteria			R
	Intelligence	Entrance	Entrance plus x (Intelligence)	
C 1-6 (1934)	.51	.49	.58 (x = 1)	.62
B 1-6	.55	.51	.62 (x = 3)	.65
L IV B	.59	.59	.71 (x = 2)	.69
L IV C	.64	.74	.79 (x = 2)	.84

In the above table R is the multiple correlation coefficient.

From this it would appear that a weighted combination of the two tests gives the most effective prediction and, from a study of the correlation scatter diagrams, it appears that this combination selects the poor pupils slightly better than the intelligence test and the best pupils slightly better than the entrance examination.

It was found that the intelligence test results correlate/

correlate less highly with the criteria at later stages in the pupils' school course. As raw scores have been used in these correlations it is suggested that this falling away might not be evident if quotients had been used. It is also suggested that temperamental factors such as industry begin to play a larger part in the child's performance as he progresses through the school.

In Germany and Sweden investigations into the prognostic value of tests and examinations have been sponsored by the International Examination Enquiry Committee.

The German report, the Schulerauslese[1] deals with

---

1. Schulerauslese. Kritik und Erfolge. O. Bobertag. Berlin, 1934.

---

examinations at different stages of the educational system. The contributions dealing with the entrance to post-primary schools were made by Dr. Bobertag in Berlin and Dr. Valentiner in Bremen.

In Germany under the Weimar Constitution, each of the secondary schools conducted an examination of its own for the selection of its pupils. This naturally gave rise to certain anomalies as a child could be rejected on one area but yet a less able child accepted in another area. The method of examination generally adopted was either an examination given by the secondary school teachers on the day of entrance or by the record/

record of school marks in the common fundamental school along with a psychological report. In the first method all the faults of a single examination were evident. To the second method, also, there were several objections. The school marks varied in standard while the psychological report was often inadequate or disregarded by the secondary schools. Conditions became critical after the Great War as the secondary schools were being flooded with pupils who did not have the necessary ability and there was a resulting overcrowding in the professions.

In 1931, a Prussian regulation stressed the tightening up of the entrance of pupils to secondary schools. The abolition of the single examination was advocated and the use of tests encouraged. Intelligence tests were little used in Germany at this time. Bobertag[1.] carried out three investigations in one of which[2.] he found (1) the differences in ability among

---

1. I. Übereinstimmung von Leherurteil und Test prüfung. Otto Bobertag, 1929.

II. Variabilität und Konstanz von Begabung und Schulleistung. Otto Bobertag, 1931.

III. Studentauslese in Amerika. Otto Bobertag, 1931.

2. In II above.

---

pupils as revealed by intelligence tests are far greater than differences in school achievement as stated in school marks,

(2)/

(2) the ordinary school reports (marks) have less prognostic value (for achievement in the secondary school) than the results of intelligence tests (3) the selection of pupils for the secondary school would have been approximately 70% better had the results of intelligence testing been taken into adequate consideration.

The Psychological Institute of the University of Munich in a report dealt with the importance of the method of selection of pupils for secondary education. In addition to school records and intelligence tests it advised that a medical report should also be taken into account.

Four means have been used for the selection of pupils for secondary schools. The fundamental teachers' estimate of fitness for secondary education, the fundamental school marks, the entrance examination and the intelligence test.

In the case of the first, these estimates were given as "highly recommended", "recommended", "not recommended" and "dissauded". Of 249 pupils in Berlin secondary schools the numbers falling into the respective categories were 77, 154, 14 and 4. Of the 77, about 80% were regularly promoted, of the second group about 50% while of the third and fourth groups together only about 11% gained promotion. The conclusion from this experiment was that "the secondary school accepts too great a number of students not sufficiently capable. This in part is/

is a result of a tendency of the common fundamental school to overrate the fitness of their pupils for secondary schooling."

To test the prognostic value of the other means of selection, two Berlin secondary schools were selected and after consideration of the pupils records it was concluded that "the final opinion of the common fundamental school, the marks and the scholastic examination are of somewhat the same prognostic value. The intelligence test has a higher prognostic value in the positive as well as the negative direction". This conclusion is similar to many of the English investigations.

An interesting point made is that the present selective method is no more effective than if no method of selection was used at all. The coefficient of alienation of the correlation coefficients in several of the investigations cited in this section lead to a similar conclusion. It must be remembered that a correlation of .80 means a relationship only 40% better than a chance one. The benefit of a uniform examination, however, it is pointed out, is that it gives an objective standard over an area.

Another conclusion is that a probationary period of eight to ten days in the secondary school would be effective in selecting pupils likely to make a success of the course. A probationary period has been suggested by several English investigators although the period suggested for probation varies from one year (Valentine) to one term (Dr. Earle).

An/

An experiment reported by Valentiner was carried out in Bremen and deals with the prognostic value of school marks, psychological reports, examination in school subjects and intelligence testing for the career in the secondary school. Each is dealt with in turn.

With regard to the marks of the common fundamental school it was found that they were generally lower than those of the secondary school. In fact school marks have "only a relative meaning"; they are fixed in relation to the achievement of the class with reference to a norm established by experience.

The psychological reports of the common fundamental school, also, had no great predictive value for,"even if the teacher is well informed concerning his pupils, his ability to give an accurate opinion of pupil fitness for secondary school work is limited, as he has no way of knowing how the personal characteristics of the child will develop under the influences of the secondary school." An experiment was conducted with the purpose of improving the predictive value of the teachers' estimates. Three things were done (1) a scientific study of their reports was made (2) special courses were given (3) the introduction of the "Bewahrungslisten" by which the teachers were given the report of the child at the end of the first year of the secondary school. From 1927 to 1932 the percentage of failures/

failures among the pupils recommended by the common fundamental school fell from 33·6% to 24·2%.

A follow-up experiment was conducted with 258 pupils during five years of their secondary course. The results were as given in Table XXXI.

TABLE XXXI.

	% failures according to			
	School Marks.	Examination in school subjects	Intelligence Testing	Combination of all
for highest Qr.	14	11	2	0
for second Qr.	24	27	16	29
for third Qr.	35	42	47	31
for fourth Qr.	63	55	69	69

From this table it is concluded that the intelligence test has the best predictive value, better even than that of the combination of the different methods.

Examinations for the purpose of selection must not be merely a means of assessing achievement but should be essentially predictive in character. The suggestion is therefore made that, as far as attainment is concerned, minima should be established below which a pupil has no hope of achieving success in a secondary school.

A special investigation was conducted on the value of the common fundamental schools psychological reports on pupils applying/

applying for admission to secondary schools. From the reports of secondary pupils in four Berlin schools 100 typical cases were selected. These were examined first of all as to their value in selection for admission to secondary schools "assuming they were correct".

The sample was examined by nine experts who were asked to mark them on a five point scale 1, 2, 3, 4, 5 with a view to their selective value. It was found that the experts differed in their making even more than was expected. In general the conclusions were that the reports were practically valueless and that teachers' marks, even though the standard does vary, are to be preferred - "measuring is absolutely necessary." One of the main reasons for the failure of the reports is the fact that the common fundamental school teachers are unfamiliar with the important aspects of secondary school work and the only chance of making the reports a success is for both the elementary and the secondary school teachers to be trained for this purpose.

Another section of the investigation dealt with the correctness of the fundamental school reports, the criterion being an estimate of fitness for secondary education by the secondary school teacher at the end of the first year. The estimates were chosen from "doubtlessly fit", "doubtful" and "unfit". For a group of 38 pupils the results were as given in/

in Table XXXII.

TABLE XXXII.

Recommendation by Fundamental School	No	Estimate of Secondary School Teachers after one year	No
Doubtlessly fit	31	Doubtlessly fit Doubtful Unfit	21 8 2
Doubtful	7	Doubtful Unfit	4 3
Unfit	0		

The conclusions drawn from these data were that the correctness of the reports is comparatively low and that the fundamental school teachers tend to overrate fitness for secondary schooling.

The qualities making for success in the secondary school were also subject to investigation. Actually the numbers dealt with and the method adopted render the results of doubtful value. The conclusions based on the frequency of the good qualities attributed to pupils graded as "doubtlessly fit", "doubtful" or "unfit" was that "the qualities must important for success in secondary school work are to be looked for on the side of intelligence not on the side of emotional life, will and character."

Another experiment was conducted with what is claimed to be the first achievement test battery in Germany. The test consisted/

consisted of six sections four of which were dealing with language, reading, vocabulary, grammar and spelling, the other two sections being on "pure numbers" and problems. Two investigations were made.

The first investigation dealt with the results of the test battery given to 1900 pupils from a typical selection of 50 fourth grade classes in the Berlin common fundamental schools. Dealing with individual results it is concluded that "traditional marking obliterates the differences between achievements of somewhat average degree," a feature not present with test results. In the case of group results it was shown that there was a considerable difference in the scholastic attainment of classes in different social class schools. This has an important implication as regards selection for secondary schools since the standard of teachers' estimates or marks will vary from school to school. "The only solution to this difficulty seems to be the introduction of a uniform achievement test for all applicants" and "a uniform examination of achievement will have complete success only if standardised tests are used." The suggestion is made, also, that teachers marks might be made comparable if they would use standardised tests on which to base their estimates.

The second investigation deals with the relation between "results of measuring and traditional marking". The data used were test results and school records for 2500 pupils from 71 fourth grade classes in one district in Berlin. It was/

was shown that school marks were influenced by the standing of the class as a whole and by the personal tendencies of the teacher to mark more or less severely. In a "good" class, achievements were marked lower by about 10% and "that a liberal marking teacher gives better marks for lower achievements than a severe marking one grants for even better ones." The position is summed up by the conclusion that "school marks do not state the real standing of different classes compared with one another."

Mention may be made of another investigation sponsored by the International Examination Enquiry Committee in Germany, to show how unsatisfactory is the entrance examination of the essay type. Sixteen essays by adult students seeking entrance to the Berlin Evening High School were marked by nine examiners. The mean correlation between the marks was ·519. This leads to the conclusion that "in the ranking of essays the influence of chance is absent as great as the influence of merit."

Even when the marking is restricted to one examiner the marking is unreliable. One examiner marked the 16 papers four times, the intervals between the markings ranging from "some days" to two weeks. The correlations between the different markings varied from ·314 to ·779, the average of all being ·575. Dealing with the "pass" or "fail" question, 7 of the 16 candidates were passed and one candidate was failed by all examiners. In eight cases, therefore, the candidate was passed by some and failed/

failed by others. "With fully half of the candidates the obtaining of the passing mark is largely chance."

In Sweden, an Examination Committee under Dr. Hanninger made an enquiry "into the value of the entrance test from the standpoint of the relation between its result and subsequent performances at school" and into "the feasibility of estimating the work of pupils otherwise than by examination."[1.]

---

1. The Entrance Examination in View of Later School Performances. Frits Wigforss. Publications of the Swedish Society for a Psychological Pedagogical Institute, Stockholm, 1937.

The enquiry was based on the results of entrance examinations at the Kalmar High School from 1929 to 1935. The main subjects of the examination were Mathematics and Swedish. A five year and a four year course were available. 518 pupils passed into the five year and 134 into the four year Realskola, elimination by failing being 17% and 27% respectively of the candidates. Actually, the effective elimination was 9% and 19% respectively as many of the candidates sat the examination again in the following year. Several candidates who would have failed were allowed to take the course as there was sufficient accommodation in the schools. Candidates, however, at the time of the examination were not sure whether they could rely on entrance despite failure and thus, data were available to indicate how pupils who had failed in the examination would have done/

done in the secondary school.

In order to investigate the prognostic value of the entrance test comparison between the entrance test and (1) class ranks (2) class marks (3) result of promotions at the end of the session were made.

In making the comparison between the entrance test results and those of the first session in the Realskola, the latter were obtained by adding together the marks recorded in the various subjects (Swedish, Mathematics, German, Religion, History, Geography, Biology) giving Swedish, German and Mathematics twice the weight of the other subjects since promotion depends largely on these subjects.

Dealing with rank correlations between entrance results and class marks in the first year of the Realskola, the average of 17 results was ·57, the coefficients ranging from ·33 to ·66. To find the correlation between the entrance test and the first year results of the five year Realskola all the annual groups were merged into one group. The correlation coefficient was ·53 $\pm$ ·03 (N = 494). With regard to the annual groups, individually the coefficients varied from ·48 to ·60, the average being ·54.

A more detailed study of the results shows that about 20% of the pupils did badly and were unable to meet the demands of the school. Again, "if 20% of the weakest at the entrance test had been eliminated then 3/5 of those eliminated would have/

have been the wrong person." The conclusion drawn is that "this would seem to show that it is rather hopeless to make the entrance test the basis of a good eliminatory process." But it must still be borne in mind that the guidance afforded by the entrance test is much better than none at all. The results of the four year Realskola are similar, the correlation coefficient being ·44 ± ·07 (N = 128).

In comparing the entrance test with the promotions at the end of the first year it was found that the number of non-promoted pupils was 27% of the whole number and, although there is a smaller number of non-promoted pupils in the higher entrance groups than in the lower ones, yet "in a 27% elimination about half of the rejected would be pupils who made good and half of the non-promoted pupils would have been admitted." In other words, passing the entrance test is a poor guarantee that the pupil will afterwards carry on his studies satisfactorily and that such a defective weeding-out instrument should be employed as little as possible, and it is therefore desirable that all who pass should be admitted, their suitability for study being gradually ascertained in the school. In this investigation also, therefore, the idea of a probationary period is suggested.

In the High School there was a number of pupils near the pass mark who were passed "with hesitation" and on investigation it was found that these were as likely to succeed as those/

those who had been passed without hesitation. By extending the comparison between the result of the entrance test and the result in the High School it is concluded that "from 30% to 40% of the failures would have made good in the first year of the High School."

An investigation was made, also, into the connection between the entrance test and High School performance during more than one year. A system was adopted whereby subsequent years were given a weighting, for example, the second year marks were given a weight of 2. The results in the case of the five year Realskola dealing only with "non-detained" pupils were as shown in Table XXXIII.

TABLE XXXIII.

	N	r	r'
One year gropp	494	0.53	0.53
Two year group	357	0.53	0.52
Three year group	258	0.46	0.45
Four year group	151	0.43	0.46
Five year group	181	0.37	0.40

r = correlation coefficient for the n year group after n years.
r' = correlation coefficient for the n year group after one year.

Whether the decrease in the correlation is due to the diminishing size of the group with a smaller scatter of marks, the diminution being caused by selection, or due to an actual decrease/

decrease it is difficult to tell.

Similar results are given in the case of the four year Realskola as shown in Table XXXIV.

TABLE XXXIV.

	N	r	r'
One year group	128	0.44	0.44
Two year group	96	0.46	0.37
Three year group	72	0.45	0.35
Four year group	49	0.42	0.32

The conclusion drawn is "that the connection persists in a surprising way and seems for this group to be just as strong after several years. This indicates that the estimation of the pupils arrived at by the teachers during the first school year seldom undergoes any radical alteration."

Another part of the investigation includes the non-promoted pupils as well as those who had been regularly promoted, a system being adopted whereby allowance is made for the year of detention. In the case of the five year Realskola the correlation between the entrance test results and the school marks is ·53 (N = 494) the corresponding figure for one year being ·52. This indicates that there is a high degree of relationship between the results in the first and fifth year of studies, a point which has been emphasised in several other investigations/

investigations. The corresponding results of the four year Realskola being ·50 and ·44 respectively. An analysis of the data showed, however, that after the first school year a general downward movement in the school marks took place.

The unsuitability of the entrance test as a means of excluding pupils unsuitable for High School studies is evident when it is recorded that of 494 pupils received into the High School after passing the test 198 or 40% were unsatisfactory. In the case of the four year Realskola, of 128 received after passing the entrance test 36 (28%) turned out to be unfit for studies in school. Again, if 40% of the candidates had been eliminated by the test then nearly two fifths of these would have been the "wrong person". On the other hand, of those who failed at the entrance test 20% would most likely have made good in the first class and 10% during the whole school time.

It is further pointed out that the distinction between "pass" and "fail" is very fine as about one quarter of the failed candidates belong to a marks group immediately below the pass mark hence it is concluded that "the requirements for a pass should be kept low."

The question of the validity of the entrance test raises the further question of the validity of the component parts of that test. In this case the entrance test comprised of tests in Mathematics and Swedish.

Dealing/

Dealing with the Mathematics section, it was found that the correlation between the entrance marks in Mathematics and the school results in Mathematics for the first year was ·47 ± ·04 (N = 494). If the method of ranking had been adopted the average for the 17 class sections would have been ·53 ± ·03. For the second year results, the coefficient was ·41 ± ·04. The coefficients for the respective years in the case of the four year Realskola were ·40 ± ·07 and ·37 ± ·09. The prognostic value of the Mathematics test with regard to the pupils' performance as a whole over the course is slight, the correlations for the five and four year courses being ·40 ± ·04 and ·40 ± ·07 respectively. After a discussion on the construction of the test, the conclusion is drawn that an explanation of this unsatisfactory position may be that the Mathematics test tested the quantity of knowledge rather than its quality; the mechanical skill of candidates in reckoning rather than their ability to solve problems independently. It is likely that the Mathematics test would acquire greater prognosticating value if mechanical skill in reckoning and the ability to solve problems received separate tests.

The Swedish test is equally unsatisfactory. The correlation between the entrance test results and those of the first and second years was respectively ·51 ± ·03 and ·47 ± ·04 in the five year Realskola, while in the four year course it remained/

remained ·44. Taking the pupils performance over the whole course, the entrance test correlated with these results to the extent of ·42 ± ·04 and ·39 ± ·08 respectively for the two courses. The correlations between the first and second year class results in Swedish in the five years course was ·70 and for the first year results in Swedish and the marks over the whole school course was ·68. The corresponding figures for the four year course were ·75 and ·76. From these data the conclusion is drawn that "these high values prove that it should be possible to make the entrance test more valuable for prognostication than it has hitherto been."

If the relationship between the entrance test and the first year results in Swedish and Mathematics was high then much would be gained for the correlation between the latter and the combined mark (a) for all subjects was ·92 ± ·01 and (b) for all <u>other</u> subjects was ·75. For the four year course, the respective figures were ·93 ± ·01 and ·75 ± ·04. It is important, also, that the reliability of the school marks for the two courses is given by the coefficients ·80 ± ·02 and ·78 ± ·04.

Dealing with the estimates of the Primary school teachers, the main point of difficulty is recognised from the start, namely, the difference in the standards of marking. It is surprising, therefore, to find that the correlation between the Primary school estimates and the actual performance in the first/

first year was ·52 ± ·03 which is practically the same as the results for the entrance test which was ·53 ± ·03. Again, it was found that although these coefficients were of the same magnitude, the Primary school record was more serviceable in weeding out those unfit for High School studies. "In eliminating a small percentage of the "passed", the entrance test rejects unfit and fit in the proportion of 1 : 1, while the proportion according to the Primary School certificate is 3 : 2." A similar result was obtained when the complete record for the High School course was taken as the criterion. The correlation with the Primary School record is ·50 ± ·03 as compared with ·52 ± ·03, the value with the entrance test. As before, the Primary record was shown to be the better means of eliminating the unfit.

To overcome the difficulty of the different standards of marking, groups in which the pupils had had the same Primary school teacher were taken and the results worked out for each group. The average correlation between the Primary record and the first year results for 12 groups was ·66 which indicates that, if there had been a greater "homogeneity" in the school marks, the predictive value of the primary school record would have been even greater.

A similar trend was shown in the results for the four year Realskola although the numbers dealt with were so small that/

that the results had little reliability. What is necessary is that the teachers' estimates should be standardised, an aim which, to some extent, could be realised if there was a wide application in the use of standardised tests, particularly in the main subjects of Swedish and Mathematics.

Sex differences were also investigated and, as far as the forecast of fitness for High School studies was concerned there was little difference. The correlation coefficients are given in Table XXXVI.

TABLE XXXVI.

	Correlation Coefficients between entrance tests and	
	1st year results	complete school record
Boys	.53	.50
Girls	.54	.55

Girls seem to do better than boys in the entrance test and also in the High School and it is shown that, even if both are selected from the same entrance group, the girls do better than the boys. A similar result is obtained when groups from the same Primary school are considered although the results of the four year Realskola show no such sex differences. The boys were just as good as the girls both in the entrance test and in the High School.

It is finally pointed out "that it is by no means certain that/

that the results obtained in this investigation hold good for other groups of pupils at other schools."

From the investigations discussed in Section D the following conclusions may be drawn.

The intelligence test gives the best single measure for predicting success in the secondary school. The Australian investigators showed, however, that the intelligence test was of greater value in selecting the less able candidates whereas the normal tests in English and Arithmetic were better for selecting clever pupils.

Teachers' estimates fail to have high predictive value because, in the first place, there is a lack of a uniform standard of marking from school to school. This indicates the need of a uniform examination and the more frequent use of standardised tests. In the second place, there is a lack of knowledge among the primary school teachers of the secondary school requirements. There is a strong tendency, also, to overrate pupils' fitness.

The best prediction is obtained from a weighted combination of intelligence and scholastic tests.

First year marks, also, give a good indication of a pupil's future success in the course. This implies a probationary period.

E. CONCLUSIONS.

From this survey of previous follow-up investigations certain conclusions seem to have been established.

An intelligence test is an essential component in any scheme for the selection of pupils for post-primary courses of instruction.

This conclusion was reached by all the investigators who included such a test in their scheme. Even the early investigations in which the intelligence tests were not so reliable as they afterwards became, showed that the I.Q. was often a better guide than the usual entrance examination in English and Arithmetic. In the majority of experiments, the I.Q. proved to be the best single measure for predicting success in the secondary school.

Abundant evidence was produced to show that the intelligence test often revealed ability which was not evident from the results of entrance examinations. Even to a greater extent did it reveal the lack of ability which was hidden by the successful performance in examinations. Collman and Jorgensen proved this fact by showing that the examination in English and Arithmetic picked out the clever children better than the intelligence test but the superiority of the intelligence test was more marked among pupils of lower ability.

Teachers'/

Teachers' estimates have a definite value in the prediction of success in the secondary school. This value could be increased if some simple method of standardising the marks from school to school could be devised.

A distinction must be made between 'estimates' and 'forecasts'. The former is really the record of the pupil's work in the primary school; the latter is more what the primary school teacher thinks the pupil will do in the secondary school.

In the American investigations the teachers' estimates or the primary school record gave the best prediction of success in the high school. Others showed that these estimates are fairly reliable within the individual schools but the varying standards from school to school do not justify comparisons between pupils from different schools. Another cause of weakness in these estimates is that the primary school teachers do not know sufficiently well the requirements of the secondary schools.

A uniform standard throughout the schools in an area may be achieved by the greater use of standardised tests. Another method would be to scale the teachers' marks on some objective standard. Research on this point has been remarkably small and it seems that a profitable field of enquiry into the scaling of teachers' marks could be opened up. The estimates might be improved, also, by giving the primary teachers an account of their pupils progress in the secondary school.

The best prediction of success is obtained by a weighted combination of the various measures of attainment and ability in the primary school.

The dangers of using a single measure for the prediction of secondary school success, even when that measure has a reliability of ·9, has been made quite clear. This is particularly evident in dealing with borderline cases. A combination of tests helps to eliminate chance errors in any one measure.

The marks obtained in the first year of the secondary school give a fairly reliable index of later success in the course.

Several investigators pointed out that the abilities required for success in the secondary school may not be measurable by tests given in the primary school. The greatest changes, it has been proved, take place during the first year in the secondary course, hence success in the later years of a course is best determined by the performance in the first year of that course.

The natural implication of this conclusion is that there should be a probationary period of a year in the secondary school, at the end of which time a pupil's likelihood of success in the course is fairly well established.

P A R T II.

I. THE INVESTIGATION

A. OBJECT OF THE INVESTIGATION

In 1921, when the Scottish Education Department abolished the Qualifying Examination, the Education Authorities were advised to come to "some arrangement under which those who have taught and those who are to teach particular individuals shall combine in an endeavour to estimate the potentialities of the material to be handled."[1] On the whole, the Authorities have attempted to carry out this injunction and, accordingly, there has been a considerable amount of experiment into the various methods of promotion to post-primary courses of instruction. Considering that there are 35 Education Committees in Scotland, it can readily be imagined that the number of schemes which have been in operation are numerous. The ordinary examination in English and Arithmetic, teachers' estimates, mental and scholastic tests and various combinations of these measuring rods have been adopted without any general agreement as to what system gives the best predictive value. Proof of this lies in the fact that the different Education Committees are still constantly changing their schemes of promotion.

1. Scottish Education Department, Circular No. 44, Dec. 13, 1921.

The object of this investigation is to throw light upon the questions of guidance and selection that arise at the qualifying/

qualifying stage; and, in particular, to find the system of assessment of abilities and attainments which gives the most valuable guidance as to the type of post-primary course for which each pupil is fitted. To arrive at some definite conclusion necessitates the comparison between the measures made in the primary school with the subsequent success of the pupil in the post-primary course to which he has been promoted. In other words, a follow-up experiment is necessary and this part of the thesis deals with the results of such an experiment.

B. GENERAL PLAN OF THE FOLLOW-UP EXPERIMENT.

The group selected for the investigation was the pupils presented for post-primary courses of instruction in Dundee in 1935-36. These were tested by all methods of examination or assessment that are at present being used at that stage and additional data for each pupil were collected by means of special rating cards. Two dates of presentation, December and May, were in operation and thus duplicate sets of tests and examinations had to be used. The numbers in the group were distributed as follows:-

December	1130
May	2099
Total	3229

Data from the following sources were obtained.

1. Intelligence and Scholastic Tests.

It is a fundamental principle of all scientific testing that/

that, for purposes of exact comparison, the methods of administration and evaluation of tests must be the same for all examinees, for all examiners and for all different occasions. Accordingly, there must be a standardisation of (1) the administration (2) the correction and (3) the results. The tests used were the following Moray House tests prepared by Professor Godfrey H. Thomson.

	For December Group	For May Group	Duration
	Preliminary Practice Test		10 minutes
Intelligence	M.H.T. 20	M.H.T. 21	45 minutes
English	M.H.E. 7	M.H.E. 8	40 minutes
Arithmetic	M.H.A. 7	M.H.A. 8	30 minutes

Elaborate precautions were taken to ensure that the administration of the test was standardised. The tests were given in all schools on the same days and at the same time. With the co-operation of the headteachers it was arranged that 30 was the maximum number in a room and that there was a minimum of disturbance during the testing. Intervals were altered and no one was allowed to enter the rooms while the testing was in progress.

Only students of Dundee Training College and teachers who had some experience in group testing were employed in the administration of the tests. Meetings were held of all those engaged in this work and memoranda with detailed instructions for/

for the giving of the tests were issued. These instructions were rigidly adhered to by all the testers.

All the tests were returned immediately after the completion of the testing to Dundee Training College. Here, the correction of the tests was carried out by voluntary assistance from students and teachers. The correction was purely objective each question having a definite answer which received one mark if correct. To ensure accuracy all the correction was checked.

In addition to the children in the Qualifying classes all those in the age-group $11\frac{1}{2}$ - $12\frac{1}{2}$ years, irrespective of their class in school, were tested so that norms for the tests could be calculated.

2. Qualifying Examination.

The same attention to detail was paid in the administration of the Qualifying Examination. A team of teachers and students carried out the examination with detailed instructions so that the procedure was standardised throughout all the schools. The organisation for the administration for the Qualifying examination was, in fact, almost identical with that of the mental and scholastic tests. Such a procedure is unusual with examinations and might, therefore, have had some bearing on the results.

The/

The examination consisted of the following papers:-

Examination	Duration	Possible Marks.
Dictation and Spelling	(45 minutes app.)	25
Composition	35 minutes	45
English General	1 hour	80
Mental Arithmetic	10 minutes	20
Written Arithmetic	1 hour	80

These papers were made up by the Director of Education for Dundee with questions selected from Qualifying examinations set in recent years in different parts of Scotland. The English and Arithmetic examinations were set on consecutive days.

A duplicate examination was drawn up for the May group in order to investigate the reliability of the Qualifying examination. These examinations were designated Examination A and Examination B and were set within a week of each other under similar standardised conditions.

The correction of the examination was undertaken by a paid examiner who had a wide experience of correcting Qualifying examination papers in various areas. No instructions were given as to the method of correction, the allocation of the marks being entirely at his discretion. After the correction had been completed the examiner was asked to divide the pupils into four categories.

A/

A - Those fitted for a 5-year secondary education.

B - Those fitted for a 3-year secondary education.

C - Those fitted for a secondary course of less than 3 years.

D - Those not fitted for promotion to one of the above types of courses.

In the case of the duplicate examination in May, only the scripts of a representative sample of pupils were corrected. Since the size of the May group was approximately 2000 it was decided that 500 would give a sample of adequate size from which any facts concerning the total group could safely be deduced.

The sample, to be representative, had to contain pupils of all levels of attainment and have a frequency distribution of marks similar to that of Examination A. This was done by arranging in order the pupils in the May group according to their marks in Examination A and selecting every fourth name to be a member of the sample. As the selection of every fourth pupil did not quite complete the 500, the additional names were selected at random from the middle of the distribution. If any pupil selected for the sample was absent from examination B then another having the same mark in examination A was substituted.

3. Teachers' Estimates.

At the time of the investigation no Qualifying examination was set, the estimates of the teachers in the primary schools being used for this purpose. On the basis of these estimates/

estimates the Education Committee decided as to the pupils who should be promoted to post-primary schools, who should be transferred to special classes for backward children, and who should be retained in the primary schools with a view to later presentation. There was no compulsion as to the choice of course for pupils who were regarded as fit for promotion to three or five year secondary courses and the headmaster generally advised the pupil as to the type of course he should follow. Any pupil paying the necessary fees was admitted to the secondary schools, but pupils who would have been debarred by reason of the expense involved, were required to sit a special bursary examination.

The teachers' estimates of attainment were for English (including History and Geography), Arithmetic, Practical Subjects and an average mark for all subjects. Estimates for English, Arithmetic and the average for all subjects were given in the form of a percentage while that for Practical Subjects was a categoric mark (F, F.G., G. etc.)

4. Rating Cards.

These cards, a copy of which is shown on page 122 were completed by the head teachers in respect of each child in the qualifying classes.

A detailed list of directions for filling up the Rating Cards was sent to each school. The grading for health by/

INTERNATIONAL EXAMINATION ENQUIRY

RATING CARD

(1) NAME ..... DATE OF BIRTH .....

(2) SCHOOL ..... CLASS ..... SECTION .....

(3) GENERAL INTELLIGENCE I.Q .....

(4) GENERAL HEALTH	INDUSTRY	HUMANISTIC SUBJECTS (LANGUAGE, ENG., HIST.)		REALISTIC SUBJECTS (ARITH., NAT. STUDY)		PRACTICAL SUBJECTS	
		INTEREST	ATTAINMENT	INTEREST	ATTAINMENT	INTEREST	ATTAINMENT
A	A	A	A	A	A	A	A
B	B	B	B	B	B	B	B
C	C	C	C	C	C	C	C
D	D	D	D	D	D	D	D
E	E	E	E	E	E	E	E

(5) SPECIAL TALENT IN (i) ART, (ii) MUSIC, (iii) GAMES, PHYSICAL EXERCISES, ETC.,

(6) PHYSICAL DEFECTS WHICH WOULD AFFECT CHOICE OF POST-PRIMARY COURSE :-

.....

(7). MARKED PERSONAL QUALITIES .....

.....

(8) HOME CONDITIONS WHICH MIGHT AFFECT CHOICE OF POST-PRIMARY COURSE :-

.....

(9) PROBABLE SUCCESS IN

FULL LEAVING CERTIFICATE COURSE	3-YEAR TECHNICAL COURSE	3-YEAR LITERARY COURSE	3-YEAR COMMERCIAL COURSE
A	A	A	A
B	B	B	B
C	C	C	C
D	D	D	D
E	E	E	E

by the Schools Medical Officer and the I.Q. from the results of the testing were entered on the cards before they were issued to the schools.

The gradings were indicated by drawing a ring round the letter corresponding to the category in which the pupil was placed, e.g. (C) . The significance of the letters were given as follows:-

A - Exceptionally high

B - High - decidedly above average

C - Average

D - Low - decidedly below average

E - Exceptionally low

In addition, it was pointed out that in a normal class the percentage of pupils falling into the various classes would be as follows:-

Category	Percentage
A	3 - 4
B	24
C	45
D	24
E	3 - 4

No account of age was to be taken in these gradings. In the case of the estimates of the probable success in the various courses, teachers were advised to make these with the greatest care in the light of all the data on the card. In particular/

particular, account was to be taken of the pupils general health and of any serious physical defects.

The significance of the letters in this case was given as

A - would complete the course with distinction and obtain the relevant leaving certificate - Full Leaving Certificate or Day School Certificate (Higher).

B - Would complete the course successfully and obtain the relevant leaving certificate.

C - Success doubtful.

D - Would fail to complete the course successfully and would be unable to obtain the relevant leaving certificate.

E - would be a hopeless failure in the course.

5. Results of the Bursary Examination.

As has been stated, the secondary schools in Dundee at the time of the investigation were fee-paying and there was no form of test controlling admission. Bursaries to these schools, however, were awarded on the results of a special bursary examination which was held twice a year in December and April. Pupils whose parent's income exceeded a certain amount were debarred from applying for these bursaries.

The examination consisted of two papers, set by the Director of Education, one in English and one in Arithmetic. The English paper was of $1\frac{1}{2}$ hours duration in the afternoon and the Arithmetic for $1\frac{1}{4}$ hours in the forenoon of the same day. Pupils sat the examination in a central building which was unfamiliar to them all.

The/

The results of the December (1935) and the April (1936) examinations were made available by the Dundee Education Committee. Although there were two examinations, the only entrance date for secondary schools was in August. Pupils who gained a bursary in December were transferred to a Central School till August when they started on their secondary course at the same time as the pupils who gained bursaries on the results of the April examination.

C. THE FOLLOW-UP.

The essential purpose of this investigation is to find the measure or combination of measures which gives the most accurate prediction of success in the post-primary school. Actually, the first step must be the definition of school success and the establishment of the means by which this can be measured.

School success is generally synonymous with success in school examinations although, as Valentine points out, "there are other kinds of ability required and revealed not only in later life but in the life of the school, which examinations do not test". The child who has attended a secondary school carries away from that school more than a mere record of marks scored in examinations. The corporate life of the school, for example, has an effect on character, a social training and other similar intangible results. Despite this, however, there is a general/

general relationship between success in examinations and success in after life.

There are three measures which might be taken as a criterion of success in school examinations.

(1) The post-primary school term examination marks.

In Dundee each pupil in a post-primary course, with the exception of those in the special classes for backward children, has a record booklet (see specimen page in Appendix II). In these are entered the marks in every subject of the course and the average mark for all subjects for each term and each session.

(2) The award of a certificate marking the successful completion of the course.

There are three such certificates in the Scottish post-primary schools. The Day School Certificate (Lower) for a two years course; the Day School Certificate (Higher) for a three years course; and the Leaving Certificate for a five years course.

The criticism which may be levelled against this measure is that it draws a hard and fast line between success and failure. No such division exists as there is a continuous range from complete failure to outstanding success.

(3) The teachers' estimate of the pupil's success in the course.

This measure is, to some extent, bound to be subjective.

All three criteria will eventually be used but for the purposes of/

B = Total Bursary examination mark

Be = English mark in bursary examination

Ba = Arithmetic mark in bursary examination.

II. THE STANDARDISATION OF THE TESTS

A. STANDARDISATION

To be thoroughly objective a test must be standardised in its three aspects, administration, correction and results. It has been shown that the first two aspects were adequately dealt with and it is with the third that this section deals.

Standardisation entails the calculation of a line of norms for the test and the establishing of a means for the conversion of the scores into some generally accepted measure, for example, quotients. Norms may be defined as levels of attainment which represent the average type of achievement for the whole group in question. They constitute, also, the means by which the degrees of abnormality shown by examinees above and below normal, can be measured.

Now the normal achievement of a class in one school will not necessarily be the same for the corresponding class in another school. Indeed, the performance of a group in a school situated in a poor industrial area is generally lower than that of the corresponding group in a school situated in a good industrial area. The standard is not fixed but varies according to the type of pupil being tested. Generally, in any investigation, it is necessary to compare the examinees not only among themselves, but in relation to a standard, which is the/

the standard of achievement not of a section of the population but of the entire population. The ideal procedure would be to find the standards of achievement for pupils of the specific age over the whole country, as was done in the case of the Scottish Mental Survey. This, naturally, was beyond the scope of this investigation, and so the tests were standardised on the community in which the experiment was being conducted, namely, the school children of Dundee. This gave, at the same time, a comparison with any other area for which the test has been standardised. It must be remembered, however, that a pupil with a quotient of 100 on the Dundee standardisation would not necessarily have the same quotient if the standardisation on another area had been used.

The standardisation technique was that devised by Professor Godfrey H. Thomson, full details of which are given in an article in the British Journal of Educational Psychology[1].

---

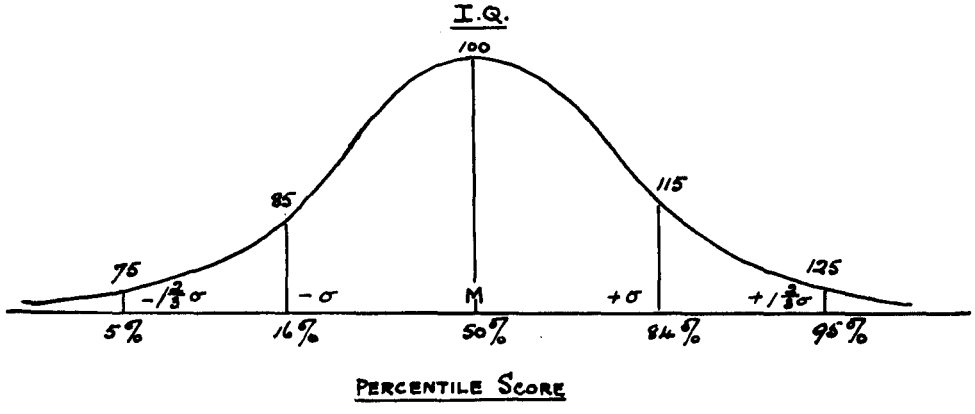
1. The Standardisation of Group Tests and the Scatter of Intelligence Quotients. A Contribution to the Theory of Examining.
G.H. Thomson. Brit. J. Educ. Psych., Vol. II, Parts I & II, 1932.

To standardise a test, the first essential is that a numerous and uncreamed age-group should be tested. From the monthly score-scatter and the normal scatter of intelligence quotients the line of norms may be determined. Thomson decided to/

to use 15 as the standard deviation of I.Q.s although he states that there is a mass of evidence to show that this figure may be more accurately taken as nearer 20. Having fixed the scatter of intelligence quotients, it is possible to arrange the line of norms so as to give this degree of scatter. The conversion of scores into quotients assumes, of course, that the test has been constructed so as to give the scatter of marks along a normal frequency curve.

FIGURE I.

Normal Distribution of Scores and Quotients.



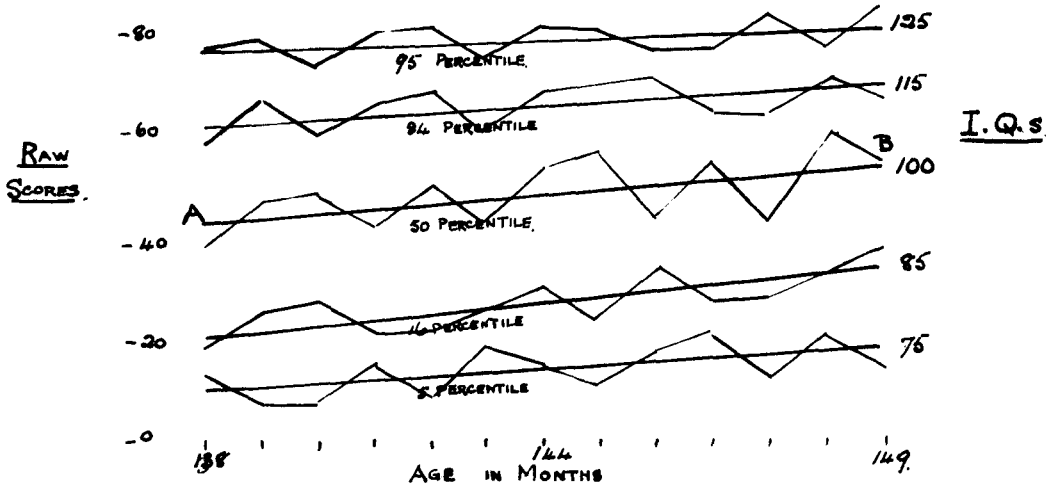
The mean score for a month is taken as corresponding to an I.Q. of 100 for that age. Since, in a normal distribution, a score of one standard deviation above the mean cuts off approximately 84% of the total group, then the 84 percentile score corresponds to the I.Q. 115 (mean + σ). (See Figure I above). Similarly the 16 percentile score corresponds to I.Q./

I.Q. 85 (mean– σ). 95% and 5% of the cases are cut off by scores of $1\frac{2}{3}$ of a standard deviation above and below the mean which will therefore correspond to I.Q.s of 125 and 75 respectively. Intermediate quotients with their corresponding scores can be found by interpolation. This can then be done for each month in the age-group.

The technique for making these calculations may be described briefly in the following manner.

FIGURE II.

Conversion of Scores and Quotients.



From the monthly distribution of scores, the zig-zag representing the monthly medians is drawn (Fig. II, AB), and to this a straight line is fitted by the method of least squares [1].

1. op. cit. Appendix I. page 135.

This/

This line is taken to represent I.Q. 100. Similarly the 84 percentile line and the 16 percentile line represent I.Q. 115 and I.Q. 85 respectively. I.Q. 125 and I.Q. 75 are represented by the 95 and 5 percentile lines. The intermediate lines may be inserted by interpolation.

In actual practice a ready reckoner is substituted for the graph. This is used by reading along the row appropriate to the pupil's age till the score is reached, when his I.Q. is found at the top or foot of the column. If the score falls between two given scores, then, with the aid of a subsidiary reckoner (vernier) an I.Q. intermediate to those at the head of the two columns is deduced.

The Moray House tests which were used in the investigation had already been standardised for an English County. At first it was thought that these standardisations might be suitable for the purpose of the inquiry, but there were two objections to this procedure. Firstly, there was some doubt as to whether the results of an English county like Northumberland would be similar to those of a Scottish industrial town such as Dundee. There was the possibility that the two age groups might be of different levels of ability, and again, the conditions under which the tests are given in England differ from those under which such an experiment is carried out in Scotland. For example, it is a general rule in England, that these/

these tests are given to determine which children will obtain free places in a secondary school, and this gives the candidates an incentive to do well. Added to this is the fact that such tests have been given for several years in the county of Northumberland with the result that the children have acquired a certain familiarity with the tests. These problems had already been experienced by the Scottish Council for Research in Education when they undertook a National Survey of a complete age group in 1932.[1] Secondly, the age groups selected for the

1. The Intelligence of Scottish Children, University of London Press, 1933. pp. 72-79.

standardisation of the tests in England were at a lower age level than that of the bulk of the group to be tested in the inquiry. The English age group was from 10 years 9 months to 11 years 8 months whereas most of the children in the Qualifying group were in the 11 years 6 months to 12 years 5 months age group. It was decided, therefore, that the tests should be standardised for Dundee.

From a previous survey of the Qualifying children in the town it was found that the average age, near the end of a session, was 12 years. About 47% of the children fell into the 11 year group and an equal percentage into the 12 year group, the remaining 6% being distributed from 10 to 14 years. To standardise/

standardise a test for the 11 year and the 12 year age groups would have involved over 6000 tests and as there were six different tests being used it meant that well over 36,000 scripts would have been necessary. Financial and administrative difficulties forced the adoption of another method of standardisation. On a further analysis of the age distribution of the Qualifying children, it was found that something approaching 73% of their number lay within the age group 11 years 6 months to 12 years 5 months. Hence, by standardising the test on this age group, and by extending the norms on either side of the age limits, a sufficiently accurate standardisation for the whole group would be established.

The age group for standardisation, then, consisted of all children whose ages lay between the limits of 11 years 6 months and 12 years 5 months, no matter the stage in any school and no matter the type of school.

As far as the intelligence tests were concerned, it mattered little whether the pupils were in primary or post-primary schools, but the effect of post-primary children on the standardisation of the English and Arithmetic tests was not quite clear. For example, in the December and May groups the distribution of the children with respect to the class in school is given in Table I.

TABLE/

TABLE I.

Class Distribution of the December and May Complete Age Groups

	Class	December		May	
		No.	%	No.	%
Primary School	Sen. III	48	1.6	45	1.5
	Sen. II	249	8.1	348	11.9
	Qual.	1860	60.3	1940	66.7
Post-Primary School	Ist Year	906	29.4	576	19.8
	IInd "	19	0.6	2	0.1
	Total	3082	100.0	2911	100.0

In the December group there are about 30% of the pupils who have had at least four or five months of post-primary instruction, while in the May group, about 20% have had similar instruction, some for four or five months, others for about a complete session. On the other hand, this may be partially balanced by the fact that a percentage of the children are in pre-qualifying classes and therefore would not have had instruction in parts of the subjects, particularly Arithmetic, knowledge of which is necessary in some of the test questions.

B. CALCULATION OF STANDARDS FOR THE TESTS.

1. Intelligence.

(a) Standards for the Age-group.

The following table gives the distribution of scores for each month of the age-group for M.H.T. 26, which had a maximum score of 100.

TABLE II.

Monthly Distribution of Scores. (M.H.T. 20)

Age	0 -9	10 -19	20 -29	30 -39	40 -49	50 -59	60 -69	70 -79	80 -89	90 -99	Total
11-6	17	27	42	42	47	27	10	12	3	0	227
11-7	21	29	42	42	47	43	21	12	3	0	260
11-8	13	24	36	51	48	39	18	11	3	1	244
11-9	9	32	28	42	49	39	20	17	0	0	236
11-10	9	30	38	44	31	45	20	16	4	0	237
11-11	11	28	33	51	46	44	20	11	5	0	249
12-0	10	12	30	44	59	40	34	16	5	0	250
12-1	8	20	38	46	45	38	28	15	6	1	245
12-2	7	9	36	47	53	41	36	15	4	0	248
12-3	6	23	26	36	49	50	24	20	9	1	244
12-4	8	15	22	42	48	31	38	18	8	0	230
12-5	12	15	25	51	50	44	31	23	10	0	261
Total	131	264	396	538	572	481	300	186	60	3	2931

For each monthly distribution the 95 percentile, the 84 percentile the 50 percentile, the 16 percentile and the 5 percentile scores were calculated and are given in Table III.

TABLE/

TABLE III.

Percentile Scores for Each Monthly Distribution. (M.H.T. 20).

Age	No.	Percentile Scores				
		5	16	50	84	95
11-6	227	6.2	16.7	36.1	55.3	72.5
11-7	260	5.7	16.6	38.6	58.2	71.2
11-8	244	8.8	20.1	39.1	58.0	72.0
11-9	236	10.4	18.5	40.9	59.3	72.6
11-10	237	10.5	19.1	39.0	60.6	74.6
11-11	249	10.0	19.7	39.8	58.7	72.8
12-0	250	11.58	25.5	44.4	63.9	74.9
12-1	245	11.7	22.5	41.8	63.4	76.0
12-2	248	15.5	26.1	44.2	63.8	73.9
12-3	244	12.2	23.4	45.9	65.8	78.4
12-4	230	11.8	25.8	45.4	66.7	77.6
12-5	261	10.2	25.4	45.0	66.7	78.2

By means of the method of least squares, the equations to the best fitting straight lines to these points (considering the percentile scores to be represented graphically) were calculated and are given in Table IV.

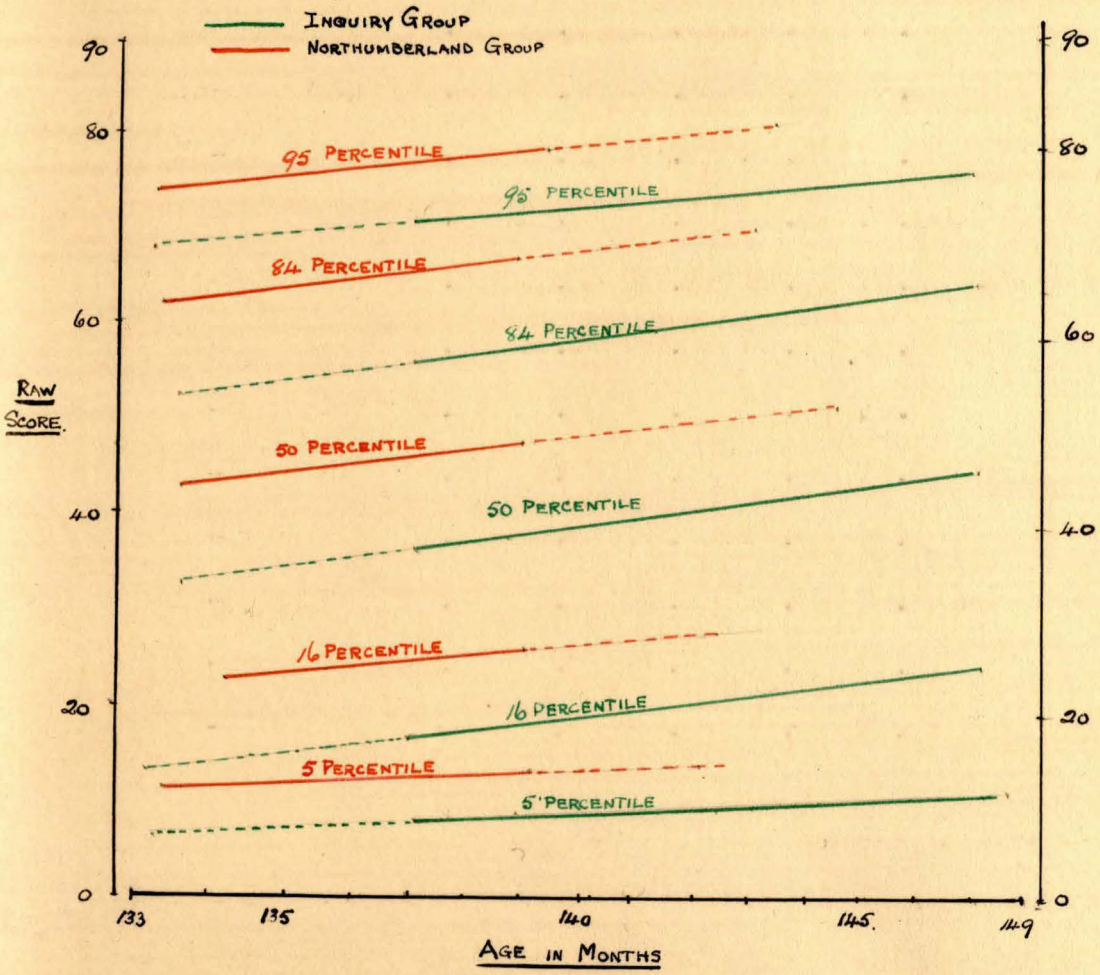
TABLE IV.

Equations to "Best Fitting" Straight Lines (M.H.T. 20).

Percentile	Equation
95	$s = .62a - 14.4$
84	$s = 1.00a - 81.8$
50	$s = .83a - 77.4$
16	$s = .90a - 107.6$
5	$s = .54a - 67.1$

where s is the score in the test and a is the age in months. These equations are represented graphically in Figure III.

FIGURE III.



The equations illustrate what Professor Thomson has termed the "ceiling effect"[1], namely that, at the higher levels of ability

1. op. cit., page 128.

a smaller age allowance per month than the average must be made due to the increased difficulty experienced by the abler children/

children in increasing their high score. At the lowest level the age allowance is also smaller than the average which may be explained by the fact that it is increasingly difficult for the less able children to score fewer marks.

Similar tables for M.H.T. 21 were constructed and are given below. The maximum score in the test was 100.

TABLE V.

Monthly Distribution of Scores. (M.H.T. 21).

Age	0 -9	10 -19	20 -29	30 -39	40 -49	50 -59	60 -69	70 -79	80 -89	90 -99	Total
11-6	6	13	23	26	36	46	26	18	16	0	210
11-7	13	9	26	25	38	44	47	25	10	0	237
11-8	9	13	23	26	33	43	40	19	8	1	215
11-9	9	13	21	30	32	41	38	35	20	5	244
11-10	9	10	19	25	28	40	42	33	19	6	231
11-11	9	13	27	19	37	39	33	32	15	5	229
12-0	9	9	10	14	46	44	36	27	24	5	224
12-1	5	11	15	23	42	51	35	38	30	10	260
12-2	9	7	15	19	26	47	52	40	30	5	250
12-3	7	6	18	17	30	37	44	38	33	6	236
12-4	7	6	8	25	42	44	33	38	29	13	245
12-5	4	4	13	17	36	45	42	42	26	5	234
Total	96	114	218	266	426	521	468	385	260	61	2815

TABLE VI.

Percentile Scores for Each Monthly Distribution (M.H.T.21)

Age	No.	Percentile Scores				
		5	16	50	84	95
11-6	210	13.0	25.9	49.7	69.7	82.9
11-7	237	8.7	25.6	51.2	68.9	78.8
11-8	215	10.9	24.9	50.3	67.9	78.6
11-9	244	12.0	27.6	53.6	75.5	85.9
11-10	231	12.1	29.0	55.6	75.9	86.6
11-11	229	11.4	24.9	51.9	74.3	85.2
12-0	224	11.9	35.1	55.0	77.0	86.9
12-1	260	16.8	34.1	56.2	79.1	88.5
12-2	250	14.5	34.2	59.9	78.3	87.0
12-3	236	17.5	33.5	60.2	79.9	87.7
12-4	245	18.3	36.8	57.4	80.5	90.1
12-5	234	22.3	39.1	59.1	78.0	86.9

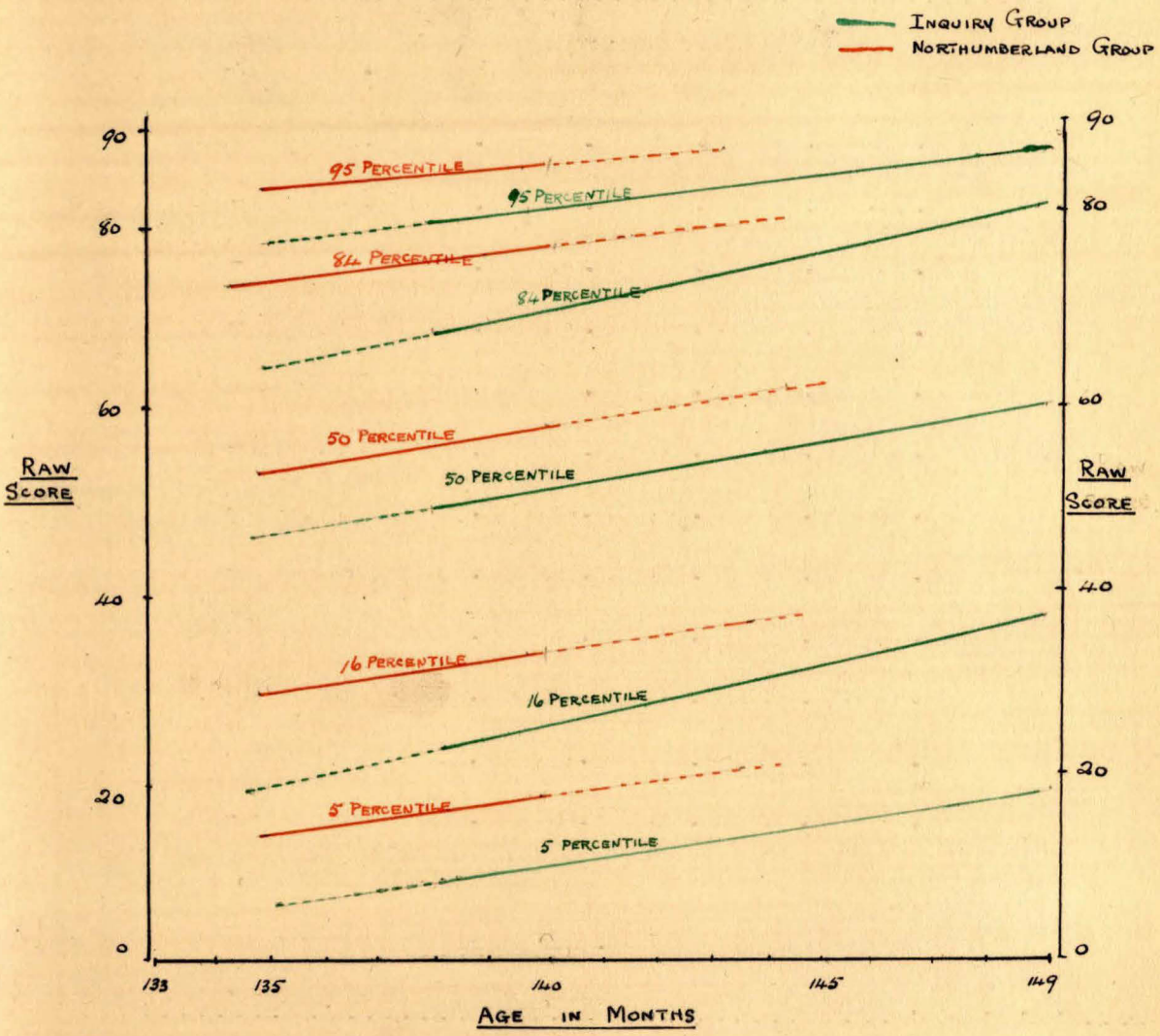
TABLE VII.

Equations to "Best Fitting" Straight Lines. (M.H.T. 21)

Percentile	Equation
95	s = .78a - 26.92
84	s = 1.07a - 78.7
50	s = .93a - 78.9
16	s = 1.28a -153.4
5	s = .92a -118.4

These equations are represented graphically in Figure II.

FIGURE IV.



The decrease in age allowance at the highest and lowest ability levels is again illustrated.

(b) Comparison with the Northumberland Standardisation.

Both intelligence tests, as has been mentioned, had previously been standardised in Northumberland on a group of

6726/

6726 children, and thus a comparison with the inquiry group was possible. The equations to the "best fitting" straight lines for the Northumberland group were for M.H.T. 20, as given in Table VIII.

TABLE VIII.

Equations to "Best Fitting" Straight Lines. (M.H.T. 20) Northumberland Group

Percentile	Equation
95	s = .72a - 21.6
84	s = .82a - 46.8
50	s = .85a - 71.0
16	s = .58a - 55.0
5	s = .42a - 45.3

Portions of these lines are represented graphically by the red lines in Figure III.

The first point to be observed is that the age groups on which the standardisation are based are at different age levels, the Northumberland group ranging from 10 years 9 months to 11 years 8 months while the inquiry group range from 11 years 6 months to 12 years 5 months. In Figures III and IV the percentile lines for each group have been produced by dotted lines so as to give a comparison over the same range of age.
The most outstanding feature of the first diagram is the superiority of the Northumberland group at every level of ability. Another feature is the parallelism of the corresponding lines, meaning/

meaning that the age allowance at the different levels are in close agreement. When translated into terms of I.Q., the difference between the two standardisations amounts to something in the region of 7 to 8 points of I.Q.

In the case of M.H.T. 21, the standardisation made previous to that of the inquiry was for the county of Northumberland. The equations to the "best fitting" straight lines are given in Table IX.

TABLE IX.

Equations to "Best Fitting" Straight Lines. (M.H.T. 21) Northumberland Group

Percentile	Equation
95	s = .67a - 6.7
84	s = .83a - 38.1
50	s = .93a - 71.6
16	s = 1.07a -120.1
5	s = .90a -107.1

These lines are represented graphically by the red lines in Figure IV, the lines being produced to give a comparison with the inquiry group. The results are similar to those of M.H.T.20 though the difference is not so marked at the higher levels of ability.

At first sight it would appear that the Northumberland children are superior in intelligence to those of the inquiry group, but there are several important points which must be considered before any such conclusion can be drawn.

In/

In the first place, as has already been mentioned, consideration must be taken of the fact that the children in the county of Northumberland are familiar with such tests since they are used anually for the selection of pupils for free places in secondary schools and this fact provides these children with an additional incentive. In the report of the National Survey of Scottish children, it has been pointed out that teachers in such a county might, in their ordinary work, give questions of a similar nature, requiring similar methods of underlining and filling in blanks. This might make the children familiar with the type of response required. The report concludes that "we are naturally inclined to hope that the poorer performance in Scotland at all levels of ability are sufficiently explained by the less intense motivation in Scotland (where nothing depended on the test) and the total absence in most districts of previous experience of such tests, or of coaching for them."[1.]

---

1. The Intelligence of Scottish Children, University of London Press, 1933, page 74.

Secondly, the standardisations were made on areas which differed to some extent from each other, the one a rural county, the other an industrial town. A later investigation with these tests showed that when the test was standardised on Newcastle children, there was an inferiority of $2\frac{1}{2}$ to 3 points of I.Q. compared with the Northumberland standardisation, a difference which the investigators consider to be "readily explained by the greater/

greater familiarity of Northumberland with intelligence tests." It would appear therefore that motivation would perhaps account for 3 points of I.Q. but an explanation has to be found for still another 3 points of I.Q.

A third point of difference lies in the age groups on which the standardisation was made. The Northumberland group had an age range of 10 years 9 months to 11 years 8 months while the inquiry age group ranged from 11 years 6 months to 12 years 5 months and, for the purposes of comparison, extrapolations had to be made, there being only a three monthly range over which the age-groups coincided. The further the percentile lines are produced the more unreliable the results become, and thus at the higher ages of the inquiry group the comparison is bound to be of doubtful validity. For the common age-range, namely, 11 years 6 months to 11 years 8 months, a comparison between the corresponding percentiles for each monthly distribution is given in the following tables.

TABLE/

TABLE X.

Comparison between Northumberland and Inquiry Group Over Common Age Range. (M.H.T. 20).

Percentile	M.H.T. 20					
	11 6/12 yrs.		11 7/12 yrs.		11 8/12 yrs.	
	I	N	I	N	I	N
95	72.5	76.4	71.2	78.3	72.0	78.1
84	55.3	66.2	58.2	67.5	58.0	69.0
50	36.1	47.6	38.6	48.7	39.1	47.4
16	16.7	25.1	16.6	27.2	20.1	26.0
5	6.2	12.6	5.7	.215	8.8	13.7

TABLE XI.

Comparison between Northumberland and Inquiry Group Over Common Age Range. (M.H.T. 21).

Percentile	M.H.T. 21					
	11 6/12 yrs.		11 7/12 yrs.		11 8/12 yrs.	
	I	N	I	N	I	N
95	82.9	85.9	78.8	87.7	78.6	86.5
84	69.7	76.1	68.9	77.5	67.9	77.4
50	49.7	56.3	51.2	56.4	50.3	58.9
16	25.9	31.4	25.6	32.9	24.9	37.8
5	13.0	15.5	8.7	17.7	10.9	18.9

From/

From the above tables, it is evident that there is a definite superiority in the ability of the Northumberland children to score more highly in these tests. That this ability may not be entirely due to superior intelligence has already been indicated.

In 1935, when the majority of the Qualifying group were in class Sen. III, a survey of all the Sen. III children in the town was carried out, the test used being the Simplex Junior Scale.[1] It was found that the mean I.Q. of the group

---

1. The Simplex Junior Intelligence Scale by C. A. Richardson. Harrap & Co. Ltd.

---

numbering 2887 was 91.6. In this case the published norms were used to convert the raw scores into quotients. To test the reliability of this result, 48 cases, selected at random, were tested with an individual Binet test. The correlations between the two sets of results was 0.75 and the mean I.Q. by the Simplex test was 96.4 with a standard deviation of 17.6 while the corresponding results by the Binet test were 108.2 and 10.5. These data seem to indicate again the apparent inability of the inquiry group to achieve normal scores in group intelligence tests.

II. English.

(a) Standards for the Age-group.

The English tests M.H.E. 7 and M.H.E. 8 were standardised with the same technique as that used for the intelligence tests.

The/

The following tables give the results for M.H.E. 7.

TABLE XII.

Monthly distribution of Scores. (M.H.E. 7).

Age	0 -9	10 -19	20 -29	30 -39	40 -49	50 -59	60 -69	70 -79	80 -89	90 -99	100 -109	110 -119	120 -129	Total
11-6	8	3	4	15	16	22	38	44	29	23	17	8	2	229
11-7	6	7	4	16	20	25	34	29	46	35	23	14	1	260
11-8	6	1	5	12	16	25	25	39	48	37	18	10	4	246
11-9	3	2	7	8	20	23	28	32	38	31	21	20	2	235
11-10	2	2	4	15	19	19	33	32	34	38	18	16	4	236
11-11	4	2	2	9	22	28	22	42	46	33	22	14	1	247
12-0	7	1	3	3	14	14	29	47	36	42	34	18	3	251
12-1	5	3	4	6	15	21	31	40	40	30	31	16	6	248
12-2	2	1	2	5	7	24	29	33	48	41	36	17	2	247
12-3	4	1	5	7	13	19	22	38	39	42	28	18	10	246
12-4	3	2	2	3	6	20	24	39	29	34	35	21	8	226
12-5	4	2	0	4	12	12	32	33	47	39	36	26	5	252
Total	54	27	42	103	180	252	347	448	480	425	319	198	48	2923

As in the case of the intelligence the 95, 84, 50, 16 and 5 percentiles were calculated and are given below.

TABLE XIII/

TABLE XIII.

Percentile Scores for Each Monthly Distribution. (M.H.E. 7).

Age	No.	Percentile Scores				
		5	16	50	84	95
11-6	229	20.8	43.6	71.4	95.3	108.6
11-7	260	19.5	43.8	75.5	98.4	110.9
11-8	246	29.8	49.1	78.0	97.5	111.2
11-9	235	29.2	48.3	77.8	102.1	114.7
11-10	236	32.0	47.3	77.0	99.6	114.6
11-11	247	34.4	49.7	77.7	98.7	111.4
12-0	251	34.8	58.2	81.6	103.9	114.2
12-1	248	30.2	52.6	79.3	103.8	115.5
12-2	247	42.9	58.9	83.8	103.8	113.4
12-3	246	32.8	54.4	83.1	105.4	118.2
12-4	226	41.7	59.6	84.3	107.4	117.9

These points represented graphically gave a zig-zag to which were drawn the best fitting straight lines, the equations of these lines being in the following table.

TABLE XIV.

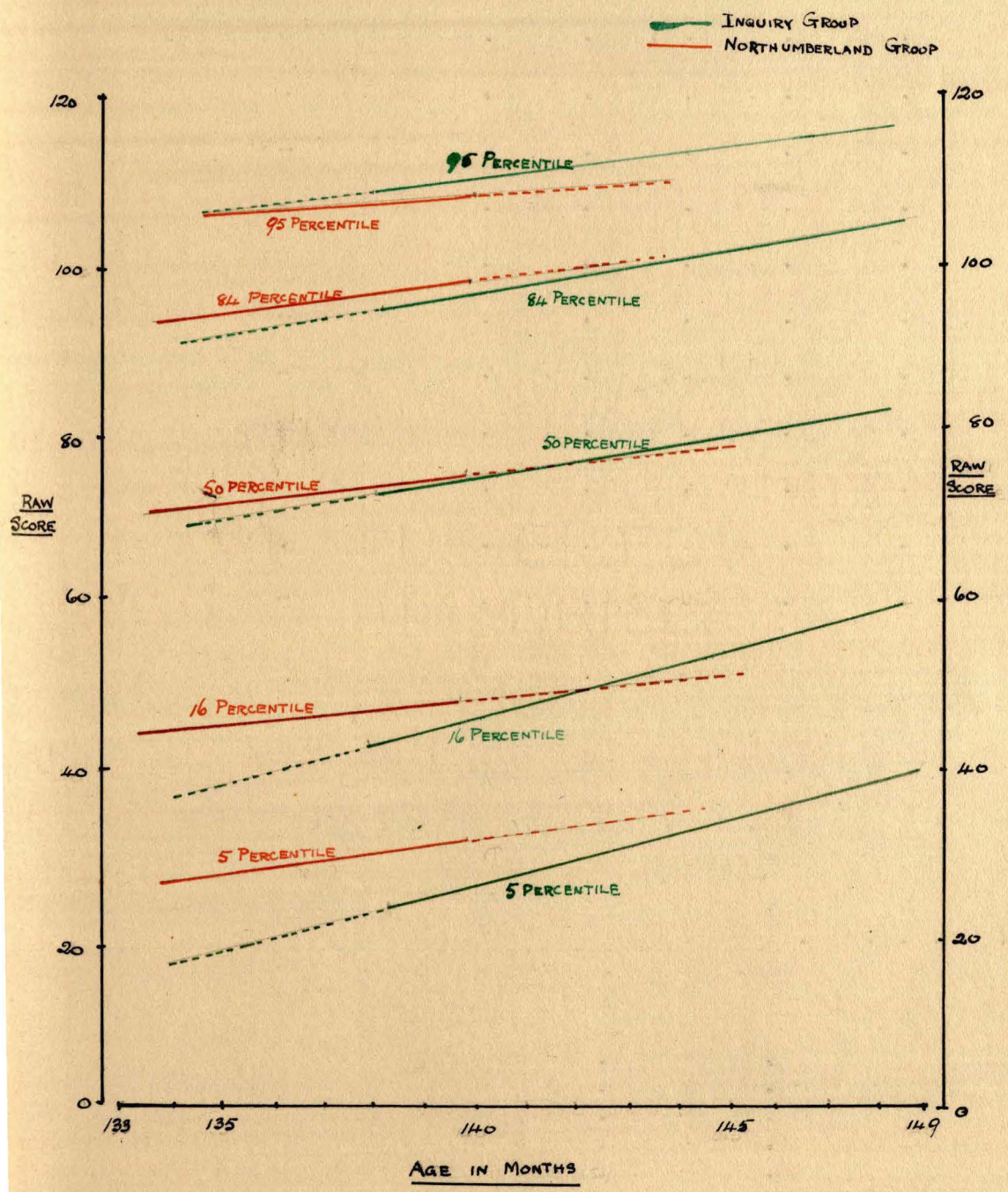
Equations to "Best Fitting" Straight Lines. (M.H.E. 7).

Percentile	Equation
95	$s = .70a + 13.5$
84	$s = 1.02a - 44.5$
50	$s = 1.07a - 74.0$
16	$s = 1.59a - 175.9$
5	$s = 1.77a - 221.5$

These/

These results are represented graphically in Figure V.

FIGURE V.



Similar results for M.H.E. 8 are given in Tables XV, XVI & XVII.

TABLE XV.

Monthly Distribution of Scores. (M.H.E. 8).

Age	0 -9	10 -19	20 -29	30 -39	40 -49	50 -59	60 -69	70 -79	80 -89	90 -99	100 -109	110 -119	120 -129	130 -139	140 -149	Total
11-6	2	7	8	15	21	20	34	26	22	22	14	10	5	1	0	207
11-7	5	8	5	13	26	31	33	34	32	26	10	14	0	0	0	237
11-8	4	5	5	14	29	25	25	37	20	27	9	9	2	0	0	211
11-9	5	8	9	12	21	28	33	29	26	28	22	16	8	2	0	247
11-10	6	5	8	8	18	19	24	36	27	26	26	11	7	4	0	225
11-11	7	4	7	15	14	24	26	28	30	28	20	17	2	3	1	226
12-0	5	2	6	11	18	19	31	36	29	26	21	10	8	2	0	224
12-1	6	3	6	10	16	21	24	44	34	37	26	16	8	3	0	254
12-2	7	3	4	7	20	18	23	28	43	46	22	13	10	4	0	248
12-3	6	1	3	10	15	23	22	33	32	31	23	21	13	2	0	235
12-4	1	3	7	6	16	25	29	32	26	37	30	14	12	3	1	242
12-5	3	1	4	4	20	24	22	37	35	27	31	21	9	1	0	239
Total	57	50	72	125	234	277	326	400	356	361	254	172	84	25	2	2795

TABLE XVI.

Percentile Scores for Each Monthly Distribution. (M.H.E. 8).

Age	No.	Percentile Scores				
		5	16	50	84	95
11-6	207	21.2	40.0	68.5	98.1	115.2
11-7	237	18.1	42.2	68.7	94.2	111.0
11-8	211	22.6	41.5	68.9	94.4	110.0
11-9	247	18.7	42.1	72.1	103.4	118.0
11-10	225	19.8	44.5	76.3	104.1	119.3
11-11	226	19.9	41.8	75.2	102.9	116.4
12-0	224	26.5	46.1	75.1	101.9	118.3
12-1	254	25.7	49.2	78.8	104.3	118.4
12-2	248	25.5	48.9	82.8	103.7	121.1
12-3	235	31.3	50.6	80.9	108.8	122.0
12-4	242	31.3	51.8	80.3	106.6	122.8
12-5	239	39.4	52.1	80.8	107.2	118.6

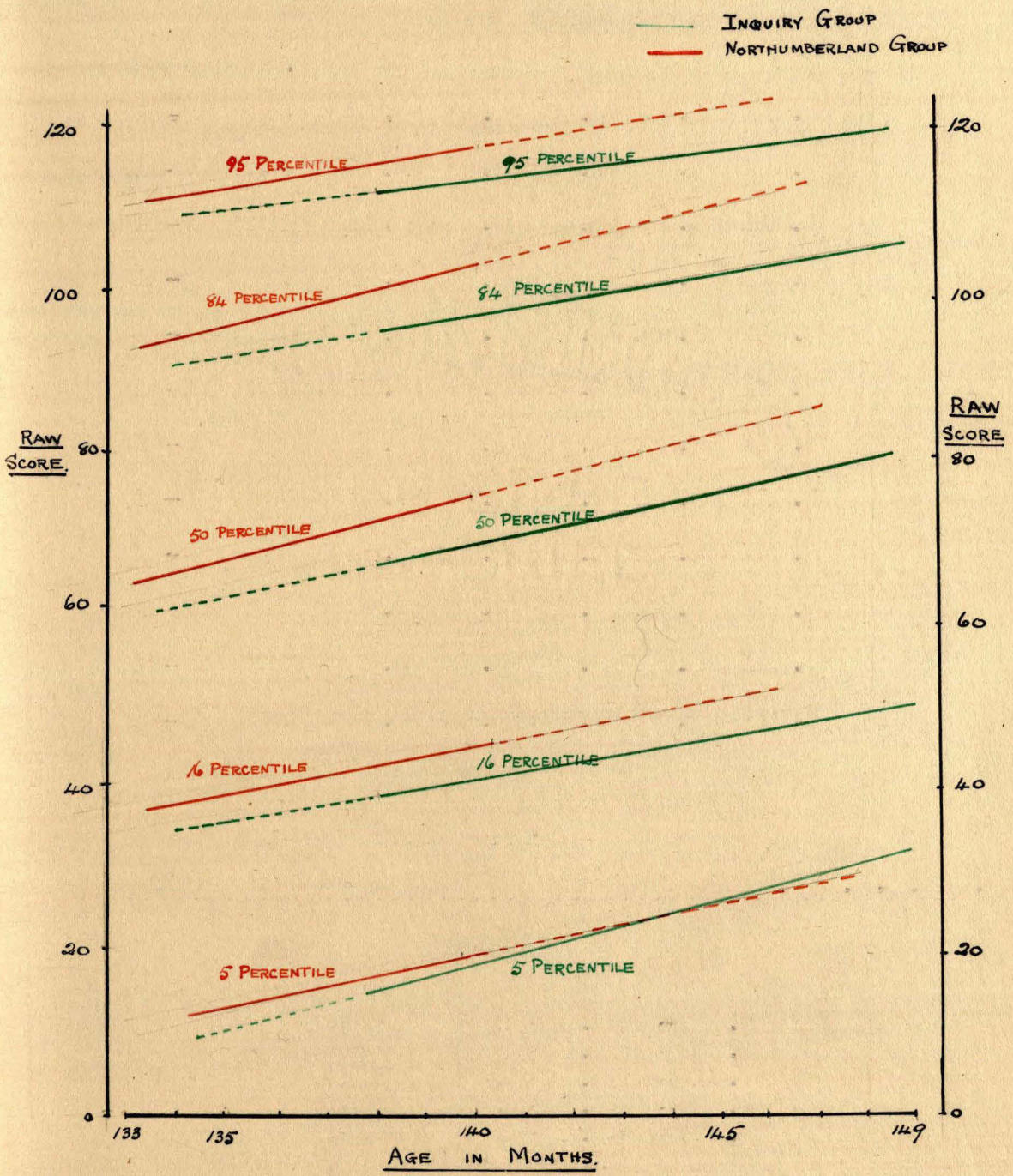
TABLE XVII.

Equations to "Best Fitting" Straight Lines. (M.H.E. 8).

Percentile	Equation
95	s = .85a - 4.8
84	s = 1.10a - 55.9
50	s = 1.34a -117.3
16	s = 1.18a -124.0
5	s = 1.53a -195.3

These results are represented graphically in Figure VI.

FIGURE VI.


INQUIRY GROUP
NORTHUMBERLAND GROUP
120
100
80
60
40
20
0
RAW SCORE.
95 PERCENTILE
95 PERCENTILE
84 PERCENTILE
84 PERCENTILE
50 PERCENTILE
50 PERCENTILE
16 PERCENTILE
16 PERCENTILE
5 PERCENTILE
5 PERCENTILE
133
135
140
145
149
AGE IN MONTHS.

These data show that the two English tests give approximately the same results. From the equations the "ceiling effect" is quite evident at the highest level of attainment, but different conditions are present at the lowest level, for, in each case, the age-allowance is greatest at this point. A possible explanation of this fact is that there are a good few easy questions in the test and this makes the scatter of marks at the lower levels of attainment greater than at any other level.

(b) Comparison with the Northumberland Group.

The equations of the corresponding 'best fitting' lines for the Northumberland results are as follows:-

TABLE XVIII.

Equations to "Best Fitting" Straight Lines (M.H.E. 7) - Northumberland Group

Percentile	Equation
95	s = .52a + 36.3
84	s = .83a - 17.3
50	s = 1.03a - 67.
16	s = .72a - 51.4
5	s = .93a - 98.

TABLE XIX.

Equations to "Best Fitting" Straight Lines (M.H.E. 8) - Northumberland Group

Percentile	Equation
95	s = 1.13a - 39.8
84	s = 1.40a - 93
50	s = 1.51a - 135.8
16	s = 1.18a - 119.9
5	s = 1.16a - 133.7

A comparison between the corresponding sets of equations is best made by reference to Figure V and Figure VI where the Northumberland 'best fitting' straight lines are drawn in red and are produced so as to compare with the inquiry group over the same range of age. In M.H.E. 7 (Fig. V) there seems to be a superiority, on the part of the inquiry group, at the higher ages and the higher levels of ability. In the M.H.E. 8 results (Fig. VI) however, there is a correspondence between the lower levels, while all others show a superiority on the part of the Northumberland group.

The comparison is more accurately determined by considering the percentiles at the ages common to both standardisations which are given in the following tables.

TABLE XX.

Comparison Between Northumberland and Inquiry Groups Over Common Age Range. (M.H.E. 7).

Percentile	11 6/12 yrs.		11 7/12 yrs.		11 8/12 yrs.	
	I	N	I	N	I	N
95	108.6	107.7	110.9	109.4	111.2	107.7
84	95.3	97.2	98.4	99.9	97.5	97.5
50	71.4	77.1	75.8	76.1	78.0	73.5
16	43.6	48.6	43.8	50.9	49.1	48.6
5	20.8	31.8	19.5	33.6	29.8	31.6

TABLE/

TABLE XXI.

Comparison Between Northumberland and Inquiry Groups Over Common Age Range. (M.H.E. 8).

Percentile	11 6/12 yrs.		11 7/12 yrs.		11 8/12 yrs.	
	I	N	I	N	I	N
95	115.2	115.7	111.0	120.4	110.0	118.6
84	98.1	99.8	94.2	102.6	94.4	103.9
50	68.5	71.5	68.7	72.2	68.9	77.9
16	40.0	40.8	42.2	44.5	41.5	47.3
5	21.2	22.8	18.1	28.8	22.6	30.1

Considering M.H.E. 7, the table shows that at the highest ability level the Inquiry group is superior to the Northumberland group but as the lower levels are approached the reverse occurs. At the lowest level the Northumberland group is definitely superior to the Inquiry group. It is interesting to note, by considering the equations to the best fitting straight lines or Figures V and VI, that the age allowance per month for the Inquiry group is greater at each ability level than for the Northumberland group. If the Inquiry group is inferior in intelligence to the Northumberland group, as indicated by M.H.T. 20, then a greater or even equal attainment in the English test would mean a greater achievement on the part of the Inquiry group.

With/

With M.H.E. 8, the Northumberland group is superior at all levels to the Inquiry group and, as regards age-allowance per month, it is higher for the Northumberland group at all levels with the exception of the lowest. Referring back to a comparison of the intelligence distribution of the two groups, it is evident that the differences in attainment of the two groups correspond to some extent to the differences in mental ability or that the reasons for the differences are similar in each test.

III. Arithmetic/

III. Arithmetic.

(a) Standards for the Age-group.

The raw scores of the Arithmetic tests, M.H.A. 7 and M.H.A. 8, were dealt with after the manner of the Intelligence and English test results. Tables XXII, XXIII and XXIV give the results of M.H.A. 7.

TABLE XXII.

Monthly Distributionof Scores. (M.H.A. 7).

Age	0 -9	10 -19	20 -29	30 -39	40 -39	50 -59	60 -69	70 -79	80 -89	90 -99	Total
11-6	20	23	39	43	53	30	19	2	1	0	230
11-7	12	33	37	52	56	40	27	5	1	0	263
11-8	12	26	28	51	63	39	20	4	1	0	244
11-9	9	34	23	35	49	51	27	5	2	0	235
11-10	8	24	44	40	56	34	19	9	0	0	234
11-11	9	24	39	50	65	43	15	2	0	0	247
12-0	7	16	37	47	62	49	28	4	3	1	254
12-1	7	28	23	41	52	55	23	5	1	0	235
12-2	5	26	32	49	59	56	21	2	1	0	251
12-3	8	21	27	48	62	43	24	7	3	0	243
12-4	7	14	28	41	46	60	27	4	2	0	229
12-5	7	19	28	56	66	48	31	4	1	0	260
Total	111	288	385	553	689	548	281	53	16	1	2925

TABLE XXIII.

Percentile Scores for Each Monthly Distribution. (M.H.A. 7).

Age	No.	Percentile Scores				
		5	16	50	84	95
11-6	230	5.3	16.8	37.2	54.6	65.0
11-7	263	9.9	18.6	39.0	57.2	66.9
11-8	244	9.6	19.9	40.3	55.9	65.9
11-9	235	10.3	17.9	42.9	58.8	67.8
11-10	234	11.0	20.7	39.7	56.7	68.1
11-11	247	10.9	21.2	39.7	54.3	62.6
12-0	254	13.1	24.3	42.7	58.6	67.8
12-1	235	11.2	20.6	43.1	57.9	67.0
12-2	251	12.4	22.4	41.8	56.6	65.0
12-3	243	11.5	23.2	42.3	58.4	68.6
12-4	229	12.7	25.1	44.8	58.9	67.5
12-5	260	12.7	25.1	42.5	58.3	66.9

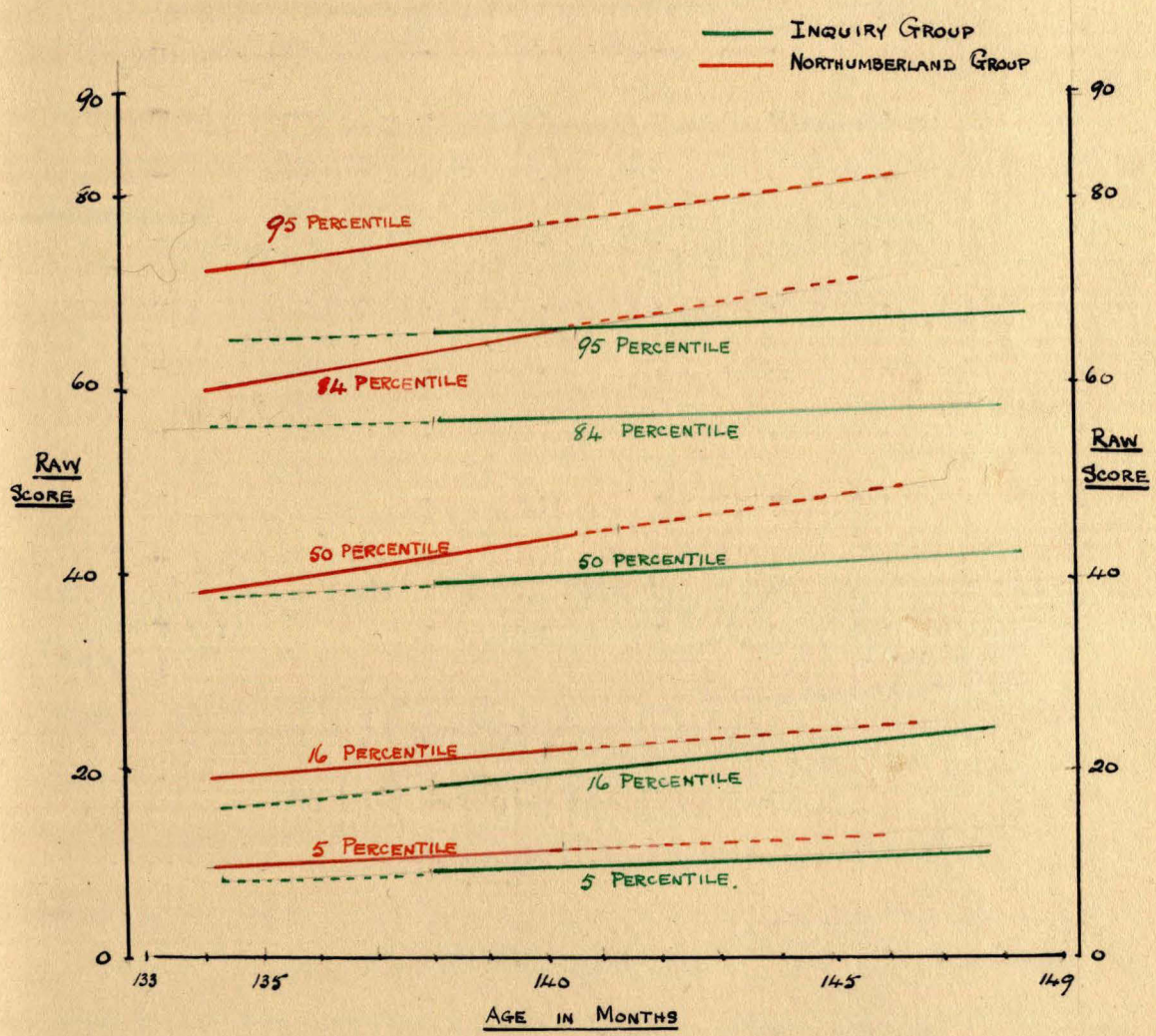
TABLE XXIV.

Equations to "Best Fitting" Straight Lines. (M.H.A. 7).

Percentile	Equation
95	$s = .12a + 49.4$
84	$s = .25a + 21.3$
50	$s = .46a - 24.7$
16	$s = .69a - 77.7$
5	$s = .47a - 56.0$

These results are represented graphically in Figure VII.

FIGURE VII.



Tables XXV, XXVI and XXVII give the corresponding results for M.H.A., 8.

TABLE XXV.

Monthly Distribution of Scores. (M.H.A. 8).

Age	0 -9	10 -19	20 -29	30 -39	40 -49	50 -59	60 -69	70 -79	80 -89	90 -99	Total
11-6	8	25	49	59	30	20	18	0	0	0	209
11-7	10	24	50	59	51	29	9	1	0	0	233
11-8	6	27	46	46	49	33	8	0	0	0	215
11-9	14	23	52	45	46	34	20	5	0	0	239
11-10	10	22	41	43	49	34	13	2	3	0	217
11-11	8	37	37	42	52	32	13	4	1	0	226
12-0	10	21	28	53	48	40	21	4	0	0	225
12-1	6	27	39	52	55	49	25	4	0	0	257
12-2	11	23	28	51	56	53	19	3	0	0	244
12-3	13	21	23	49	52	56	14	2	1	0	231
12-4	3	22	30	56	57	45	16	6	0	0	235
12-5	6	14	31	72	63	38	11	0	0	0	235
Total	105	286	454	627	608	463	187	31	5	0	2766

TABLE XXVI.

Percentile Scores for Each Monthly Distribution (M.H.A. 8).

Age	No.	Percentile Scores				
		5	16	50	84	95
11-6	209	10.5	19.6	33.3	51.8	63.7
11-7	233	10.2	20.2	35.0	50.1	58.9
11-8	215	11.3	19.8	35.7	51.5	58.7
11-9	239	8.1	19.7	36.3	55.6	66.1
11-10	217	9.9	20.2	37.8	54.6	65.0
11-11	226	10.4	17.1	36.9	53.8	64.7
12-0	225	10.1	21.3	39.6	56.8	66.1
12-1	257	12.1	21.6	40.3	57.0	66.0
12-2	244	10.0	21.3	41.1	56.3	64.7
12-3	231	8.4	20.8	41.3	55.9	63.4
12-4	235	13.5	23.7	40.6	56.0	65.9
12-5	235	13.6	25.2	38.7	52.5	59.3

TABLE XXVII.

Equations to "Best Fitting" Straight Lines. (M.H.A. 8).

Percentile	Equation
95	$s = .16a + 40.5$
84	$s = .37a + 1.04$
50	$s = .64a - 54.1$
16	$s = .41a - 38.2$
5	$s = .21a - 29.5$

certain uniform standard of achievement and there is little to differentiate the one from the other, due perhaps to the fact that so much time is spent in bringing the duller children up to a certain level that the clever children are not given the opportunity to develop their abilities to the full extent.

(b) Comparison with the Northumberland Group.

The equations of the best fitting straight lines for the Northumberland group are given in the following tables.

TABLE XXVIII.

Equations to "Best Fitting" Straight Lines - (M.H.A. 7). - Northumberland Group

Percentile	Equation
95	s = .83a - 39.4
84	s = .94a - 66.6
50	s = 1.00a - 96.0
16	s = .54a - 54.0
5	s = .25a - 24.0

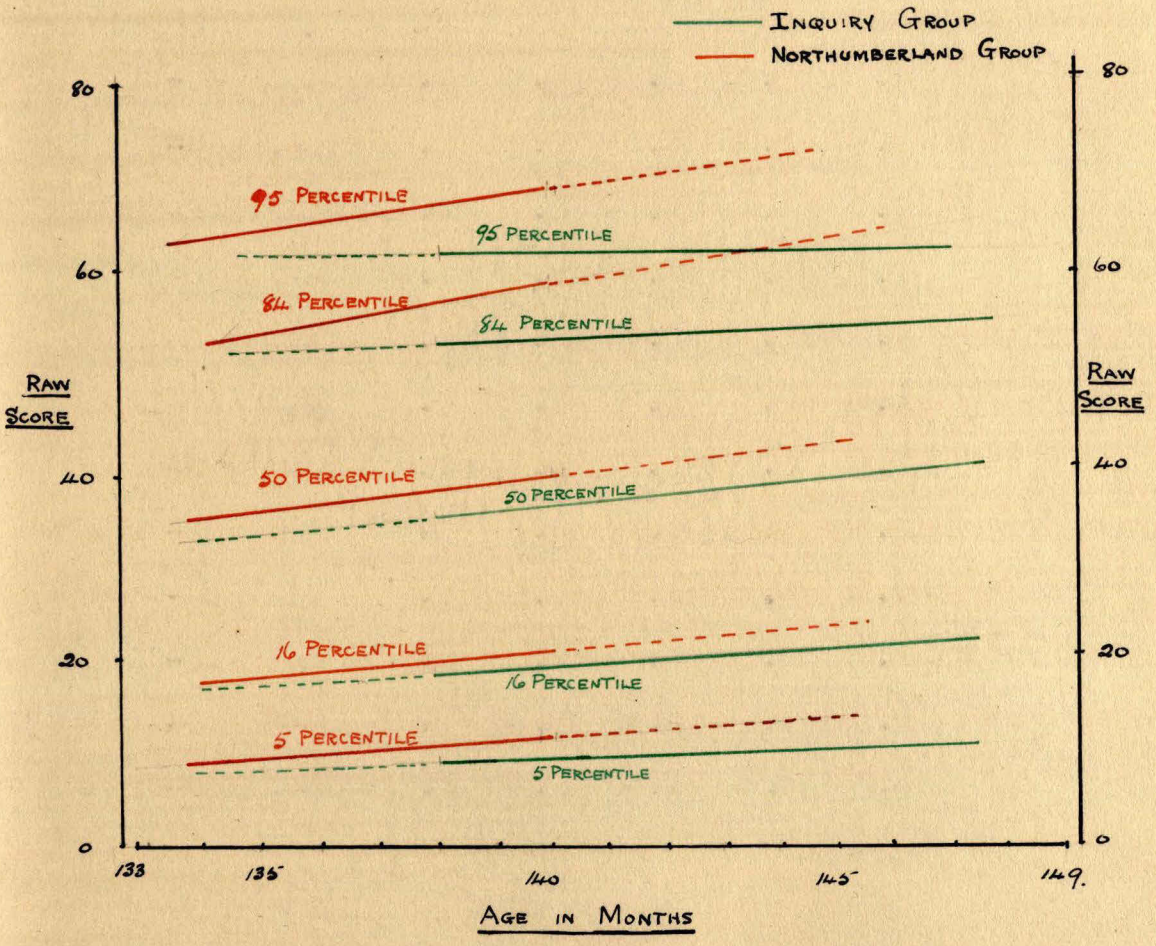
TABLE XXIX.

Equation to "Best Fitting" Straight Lines - (M.H.A. 8). - Northumberland Group

Percentile	Equation
95	s = .81a - 43.6
84	s = 1.04a - 86.7
50	s = .84a - 79.4
16	s = .55a - 55.8
5	s = .39a - 42.1

It is noticeable here also that the age-allowances per month are smaller, for the most part, than those for the corresponding English/

FIGURE VIII.



The results for the two Arithmetic tests closely correspond to each other and perhaps the most striking feature is the smallness of the age-allowance per month, particularly at the high levels of ability. The ceiling effect seems to be magnified. A possible explanation would be that the clever children reach a certain/

certain uniform standard of achievement and there is little to differentiate the one from the other, due perhaps to the fact that so much time is spent in bringing the duller children up to a certain level that the clever children are not given the opportunity to develop their abilities to the full extent.

(b) Comparison with the Northumberland Group.

The equations of the best fitting straight lines for the Northumberland group are given in the following tables.

TABLE XXVIII.

Equations to "Best Fitting" Straight Lines - (M.H.A. 7). - Northumberland Group

Percentile	Equation
95	s = .83a - 39.4
84	s = .94a - 66.6
50	s = 1.00a - 96.0
16	s = .54a - 54.0
5	s = .25a - 24.0

TABLE XXIX.

Equation to "Best Fitting" Straight Lines - (M.H.A. 8). - Northumberland Group

Percentile	Equation
95	s = .81a - 43.6
84	s = 1.04a - 86.7
50	s = .84a - 79.4
16	s = .55a - 55.8
5	s = .39a - 42.1

It is noticeable here also that the age-allowances per month are smaller, for the most part, than those for the corresponding English/

English tests. The allowance per month in the Arithmetic tests tends to be low at the low and not at the high ability levels as in the case in the Inquiry Group. Figures VII and VIII show how the results of the two tests are in close correspondence. At all levels the Northumberland norms are higher than those of the Inquiry Group. At the low levels the differences are slight but increase as the higher ability levels are approached.

The following tables give the percentiles for the common three monthly groups.

TABLE XXX.

Comparison Between Northumberland and Inquiry Groups Over Common Age Range. (M.H.A. 7).

Percentile	11 6/12 yrs.		11 7/12 yrs.		11 8/12 yrs.	
	I	N	I	N	I	N
95	65.0	75.5	66.9	76.4	65.9	77.9
84	54.6	64.7	57.2	64.8	55.9	64.9
50	37.2	44.4	39.0	44.4	40.3	43.7
16	16.8	19.2	18.6	22.2	19.9	22.0
5	5.3	9.6	9.9	10.3	9.6	10.6

TABLE XXXI.

Comparison Between Northumberland and Inquiry Groups Over Common Age Range. (M.H.A. 8).

Percentile	11 6/12 yrs.		11 7/12 yrs.		11 8/12 yrs.	
	I	N	I	N	I	N
95	63.7	67.4	58.9	68.7	58.7	71.4
84	51.8	56.4	50.1	57.1	51.5	59.5
50	33.3	37.0	35.0	36.8	35.7	40.2
16	19.6	18.8	20.2	20.8	19.8	21.5
5	10.5	11.1	10.2	10.7	11.3	12.2

In each test the Northumberland group have the better performance though less in M.H.A.8 than in M.H.A.7. At the high levels of ability the differences in achievement are very marked and here, also, the age-allowances for the Northumberland group are greater than the corresponding age-allowances for the Inquiry group. The differences become, therefore, even greater at the higher age levels which is evident from Figures VII and VIII.

III. CORRELATIONS BETWEEN THE VARIOUS MEASURES.

A. CORRELATION BETWEEN TEST SCORES.

TABLE XXXII.

Correlations Between Test Scores.

Characteristics compared	December		May		Average of December and May
	r	N	r	N	r
I v E	.891	3142	.899	3221	.895
I v A	.778	3150	.816	3175	.797
E v A	.746	3109	.769	3208	.758

The first point of note in connection with the above table is that the correlations are high relative to those obtained in many of the previous investigations mentioned in Part I. Thomson, for example, in a fairly recent article[1] and using

1. The Value of Intelligence Tests in an Examination for Selecting Pupils for Secondary Education. G.H. Thomson. Brit. J. Educ. Psych. Vol. VI. Part II. June, 1936.

Moray House tests found the correlation between Arithmetic Test scores and I.Q. to be .409 while that between English test scores and I.Q. was .438. With the data in the investigation, it makes little difference in the correlation whether scores or quotients are used.

Several suggestions may be offered as reasons for the magnitude/

magnitude of the coefficients in this experiment. The first of these is that the correlations, as far as the tests are concerned at least, are derived from large groups of unselected pupils. In many of the published investigations the population is a highly selected one and selection decreases correlation.

Another reason is that the group with which this experiment has been dealing is probably more homogeneous than that in an area which has a wide range in the type of schools. Here we have been dealing with a group of city children all of whom have been taught in fairly large schools while in other areas such as Northumberland or Fife a group may be selected from fairly large city schools down to small one or two teacher country schools.

A third reason is that, both in the case of the tests and the examination, great care was taken to ensure uniformity in administration. Experience shows that such care is especially necessary in the case of ordinary examinations which must be administered under the same rigorous conditions as intelligence tests, if the results are to be comparable.

A high correlation between the tests is not necessarily desirable especially for the purposes of prediction. If the tests correlate highly with each other and equally with the criterion it may mean that the same ability is being measured whereas the purpose in using several tests is that the range of ability measured may be as wide as possible.

B./

B. CORRELATIONS BETWEEN TESTS AND EXAMINATIONS.

TABLE XXXIII.

Correlation Between Test and Examination Results.

Characteristics compared	December		May		Average of December and May
	r	N	r	N	r
I v Qe	.822	1016	.856	1845	.839
I v Qa	.701	1017	.745	1837	.723
E v Qe	.870	1006	.861	1836	.865
A v Qa	.811	1004	.808	1827	.810

From the above table, it will be noted that Intelligence correlates more highly with English than with Arithmetic, a similar result being given in Table XXXII. The English tests, also, correlate more highly with the English examinations than the Arithmetic tests with the Arithmetic examinations. Evidence of the reliability of these results is shown in the consistency of the December and May results when different tests and examinations were used.

TABLE/

TABLE XXXIV.

Comparison Between Combined English and Arithmetic Test and Examination Results.

Characteristics compared	December		May		Average of December and May
	r	N	r	N	r
I v Q	.832	1002	.866	1809	.849
(E + A) v Q	.907	973	.894	1764	.901
I v (E + A)	.892	-	.903	-	.898

It has been seen that there is a high degree of relationship between the tests and the examinations and thus, as far as prediction is concerned, it seems immaterial which is used provided they correlate to the same extent with the criterion. The tests have the advantage over the examinations in that they are more easily administered and the scoring is much more objective than themarking of an ordinary examination. In any scheme of promotion to post-primary courses of instruction there are usually tests or examinations in English and Arithmetic. The marks in these are combined to give a total examination mark. A comparison between the combined English and Arithmetic results is given in Table XXXIV from which it seems further emphasised that the tests and examinations give similar results.

C. COMPARISON/

C. COMPARISON BETWEEN ENGLISH AND ARITHMETIC.

In Table XXXV, again, it may be noted that a constant correlation between the English and Arithmetic results is obtained whether the measures used are test scores or examination marks.

TABLE XXXV.

Comparison Between English and Arithmetic Results.

Characteristics compared	December		May		Average of December and May
	r	N	r	N	r
Qe v Qa	.685	1041	.742	1893	.713
E v A	.681	1023	.724	1901	.702

It should be pointed out that the correlation between the English and Arithmetic test scores differs from that given in Table XXXII because different groups are involved. In the case of Table XXXV only the pupils in the Qualifying classes were considered so as to make the correlation coefficient comparable to that given by a comparison of the examination marks.

D. CORRELATIONS/

D. CORRELATIONS WITH TEACHERS' ESTIMATES.

TABLE XXXVI.

Correlations Between Teachers' Estimates and Intelligence Test and Qualifying Examination Results.

Characteristics compared	December		May		Average for December and May
	r	N	r	N	r
T v I	.697	1049	.724	1941	.710
T v Q	.786	1038	.767	1848	.777

As one would expect there is a higher degree of relationship between the teachers' estimates and the Qualifying Examination marks than between the Intelligence Test scores, since the teachers' estimates are based on the results of examinations similar to the Qualifying Examination. It is surprising to find the correlations so high as the estimates have been made by quite a number of different teachers from about fifty different schools in which a uniform standard can hardly be expected.

If these estimates, however, could be scaled against some objective standard then they might be of greater value. An experiment was carried out with the results from twelve schools which supports this conclusion..

The teachers' estimates in English (Te) were compared with the English Examination marks (Qe).

TABLE/

TABLE XXXVII.

Teachers' Estimates in English Compared with Marks Scored By the Same Pupils in the Qualifying Examination.

School	Te (raw)		Qe		Te(corrected)		$r_{Qe.Te}$ (raw)
	Mean	σ	Mean	σ	Mean	σ	
1	64.7	14.7	56.6	19.9	56.2	18.9	.94
2	60.4	12.3	42.1	18.2	37.8	15.5	.72
3	64.9	10.7	62.0	17.2	60.2	15.8	.79
4	63.9	10.0	62.0	16.3	60.7	16.1	.89
5	67.5	11.1	60.7	16.7	57.8	16.9	.87
6	65.3	11.0	52.2	18.4	52.3	17.1	.82
7	74.7	10.8	58.1	18.9	54.2	18.4	.85
8	65.2	15.5	45.8	22.1	47.2	21.1	.91
9	76.1	13.2	45.9	16.5	45.8	18.1	.82
10	61.1	18.3	38.3	19.5	37.1	17.6	.83
11	60.5	7.9	39.5	15.4	36.4	12.4	.60
12	67.2	9.6	44.3	18.1	44.7	15.1	.76
Av.	66.1	13.6	51.1	20.3	49.6	19.1	.82

From Table XXXVII it is evident that the teachers' marks tend to be higher than the examination marks. The scatter of their marks, also, is much less than that of the examination which means that there is less differentiation amongst the pupils by these estimates. On the other hand, the teachers' estimates/

estimates correlate fairly highly with the examination marks within the individual schools which means that the teachers can arrange the pupils in a similar order to the examination as far as their own school is concerned.

The transformation of teachers' estimates may be simply effected if the mean and standard deviation of the estimates for each school are made the same as the mean and standard deviation of the examination marks for that school. The technique adopted was to equate the ends of the ranges and the quartiles for the two sets of marks and transform the marks from one scale to the other by means of a graph.

Table XXXVII shows how the corrected estimates correspond closely to the examination marks both in mean and scatter.

The effect of the different standards on which the estimates are based is evident in the correlations. For example when the marks for the twelve schools are slumped $r_{Qe.Te}$ is ·77 (N = 842) whereas the average correlation for the twelve schools is ·82. By using the corrected estimates the coefficient corresponding to ·77 becomes ·88.

Teachers' estimates (T) may be scaled against Qualifying examination marks or against I.Q.s. That each leads to the same result is shown by the fact that the correlation between the estimates scaled on the Qualifying Examination/

Examination marks and those scaled upon the I.Q. is ·97. Again, the correlations between Q and T for each method of transformation are

$$r_{QT\ (\text{scaled on Q})} = .88$$

$$r_{QT(\text{scaled on I.Q.})} = .86$$

$$r_{QT\ (\text{raw})} = .79$$

The results show that greater uniformity among teachers' estimates could be profitably made if they were scaled on some objective standard.

It is a difficult problem to find this standard without submitting all schools to some uniform test or examination. Once this has been done, however, there seems to be little need to scale teachers' estimates on these results.

E. CONCLUSIONS.

On the whole there is close agreement between the December and May results. The greatest difference between any two corresponding coefficients is ·057 with the coefficient $r_{Qe.Qa.}$. This difference is just significant. It may be taken, therefore, that the various intercorrelations are fairly reliable.

One interpretation of a correlation coefficient between two variables is that it is an index of the extent to which these variables/

variables measure the same ability. The intercorrelations in Table XXXII, for example, would therefore indicate that some common ability is measured by all the various tests.

Success in the English and Intelligence tests, particularly, would seem to depend upon the same ability. Indeed, the criticism raised against the English test was that, in many ways, it resembled too much a test of Intelligence. Again, success in the Intelligence test depends to a large extent on reading comprehension which is a feature of both the English test and examination. This might explain, to some extent, the high correlation between the intelligence test and the English examination.

Such an interpretation of correlation would also give some explanation of the relatively high correlations $r_{E.Qe.}$ and $r_{A.Qa.}$ and the corresponding low correlations $r_{E.A.}$ and $r_{Qe.Qa.}$, $r_{I.A.}$ and $r_{I.Qa.}$ could be explained as being relatively low since accuracy in the mechanical arithmetic section of these tests may be obtained with no great understanding of the necessary operations. $r_{Q.\ E+A}$ should, according to this explanation, be high and such is the case.

The magnitude of these intercorrelations is also important from the point of view of forming a battery of measures to predict success in the secondary schools. For this purpose, it is essential that each measure should correlate as/

as highly as possible with the criterion. If two measures, however, satisfy this condition, but correlate highly with each other, then each may be considered as measuring the same ability and hence nothing is gained by including both in a test battery. The argument for including English and Arithmetic would be stronger than for the inclusion of English and Intelligence, provided all correlated to the same extent with the criterion.

IV. THE EFFECT OF AGE ON THE CORRELATIONS

A. INTRODUCTORY

The coefficient of correlation between two variates may not only be the degree of relationship between the variates themselves but that relationship plus the degree of relationship between the variates and some other controlled factors. For example, in the case of the correlation between the Intelligence and English test scores, which was in the case of the December group $r = \cdot 891$, this coefficient may, to a greater or smaller extent, be due also to the correlation between age and the Intelligence test scores and that between age and the English test scores. In order to determine the relation between intelligence and achievement in the English test uninfluenced by the age factor, it is necessary to rule out the effect of age-differences. This may be done in two ways

(1) by the method of partial correlation

(2) by dealing with a group in which the age is constant.

(1) By the partial correlation method the coefficient of correlation between intelligence and English test scores with age constant was found from the formula

$$r_{IE.a} = \frac{r_{I.E} - r_{I.a}\, r_{E.a}}{\sqrt{(1 - r^2_{I.a})(1 - r^2_{E.a})}}$$

where/

where

$r_{I.E}$ = coefficient of correlation between I and E.

$r_{I.a}$ = coefficient of correlation between I and age.

$r_{E.a}$ = coefficient of correlation between E and age.

(2) The age range in the qualifying group is considerable as is shown in Table XXXVIII.

TABLE XXXVIII.

Age Distribution of the Qualifying Groups

Age	December 1935		May 1936	
	Number	%	Number	%
14-15	3	.3	9	.4
13-14	46	4.1	124	6.0
12-13	573	51.1	999	48.5
11-12	497	44.3	922	44.8
10-11	2	.2	6	.3
Total	1121	100	2060	100

In order to have distinct age groups and groups which were homogeneous the following analysis was made

(i) Total Group (Group T)

(ii) Qualifying Group (Group Q)

(iii) Age Group (Group A)

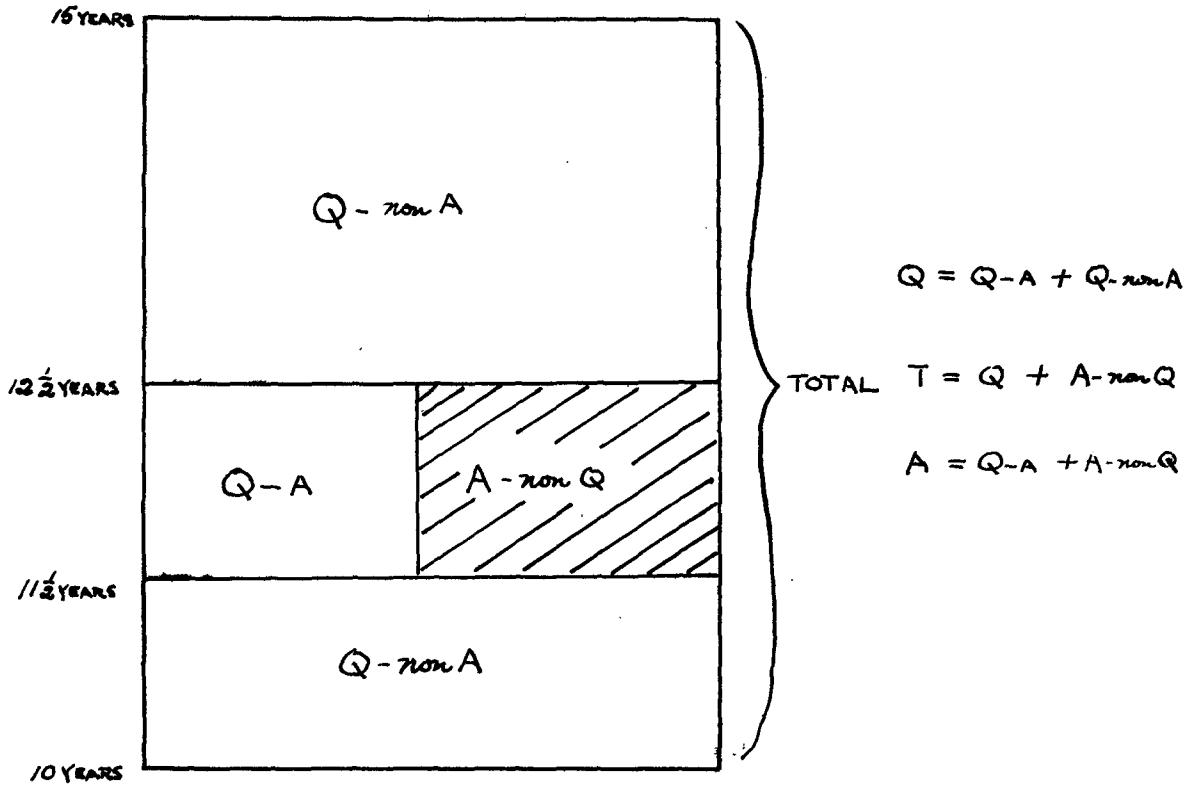
(iv) Qualifying Pupils in Age Group (Group Q-A)

(v) Qualifying Pupils not in Age Group (Group Q-non A)

(vi) Age Group not in Qualifying Classes (Group A-non Q)

This/

This may be represented diagrammatically as follows:-



The shaded portion representing the group which was not in the Qualifying Classes was composed of post-primary and Senior II or Senior III pupils. The correlations between the various measures for each group were calculated.

B/

B. CORRELATION OF TEST AND EXAMINATION RESULTS WITH AGE

The correlations between age and the scores made in the Intelligence and Scholastic tests were calculated for the groups which included all those sitting the tests (Group T). This gave a reasonable age range and a sufficiently large group. The results are given in Table XXXIX.

TABLE XXXIX.

Correlation Between Test Results and Age

Characteristics Compared	December			May		
	r	P.E	N	r	P.E	N
a v I	.060	.012	3197	.041	.012	3316
a v E	.073	.012	3162	.058	.012	3305
a v A	.030	.012	3148	.006	.012	3267

The first striking fact arising from the above table is that in the case of all the tests there is practically no correlation between the test scores and age. Such as does exist is greatest in English and least in Arithmetic. The probable errors indicate that the coefficients have practically no significance.

Table XL gives the correlation coefficients with the Qualifying Examination results.

TABLE/

TABLE XL.

Correlation Between Examination Results and Age

Characteristics Compared	December			May		
	r	P.E	N	r	P.E	N
a v Q	- .128	.023	790	- .088	.018	1404
a v Qe	- .119	.023	805	- .068	.036	1435
a v Qa	- .151	.023	789	- .091	.018	1427

From Table XL it is evident that, as in the case of the tests, there is practically no significant correlation between age and the marks in the qualifying examination. Such correlation as does exist is negative. This is according to expectation as the older pupils are generally less able, the young clever children having been promoted.

A similar result is obtained with teachers' estimates given in Table XLI.

TABLE XLI.

Correlation Between Teachers' Estimates and Age

Characteristics Compared	December		May		Average of December & May
	r	N	r	N	r
a v T	- .145	847	- .115	1535	- .130

From these results it may be concluded that the correlations given in the previous section are uninfluenced by the age factor. This conclusion is confirmed by the subsequent results.

C/

C. INTERCORRELATIONS BETWEEN TEST AND EXAMINATION RESULTS 'PARTIALLING-OUT' AGE

Table XLII shows how little change is made in the correlations when the age factor is removed by this method.

TABLE XLII.

Correlations Between the Various Measures 'Partialling-out' Age

Characteristics Compared	December		May	
	r (raw)	r (age out)	r (raw)	r (age out)
I v E	.891	.891	.899	.899
I v A	.778	.778	.816	.816
E v A	.746	.746	.769	.770
I v Q	.834	.831	.863	.865
I v Qe	.818	.815	.852	.853
I v Qa	.706	.700	.744	.745
E v Qe	.982	.982	.859	.861
A v Qa	.840	.836	.838	.839
I v T	.705	.699	.723	.726
T v Q	.797	.793	.770	.768

D. INTERCORRELATIONS BETWEEN TESTS AND EXAMINATIONS USING AGE-GROUPS

Considering the intercorrelations between the test scores there are two possible age-groups which might be selected

(1)/

(1) all pupils between the ages of $11\frac{1}{2}$ years to $12\frac{1}{2}$ years (Group A)

(2) all pupils between the ages of $11\frac{1}{2}$ years to $12\frac{1}{2}$ years who were in the Qualifying Class (Group Q-A).

The second group is a more homogeneous group than the first since all pupils in it will have reached the same stage in their primary course.

TABLE XLIII.

Correlations Between Test Scores for the Age Group

Characteristics Compared	Group T		Group A		Group Q-A	
	r	N	r	N	r	N
	December Results					
I v E	.891	3142	.892	2894	.853	779
I v A	.778	3150	.782	2901	.736	781
E v A	.746	3109	.753	2865	.689	779
	May Results					
I v E	.899	3221	.903	2737	.872	1422
I v A	.816	3175	.821	2699	.777	1400
E v A	.769	3208	.779	2727	.735	1420

Table XLIII gives the intercorrelations between the test scores for the different groups. Although, on the whole, there is little difference in the correlations there is a distinct tendency in certain directions. The correlations for Group/

Group A are all higher than the corresponding correlations for Group T. In other words, by removing the age factor by this method the correlations are slightly raised.

By a still further refinement of the groups, however, the correlations are lower than previously. One possible explanation of this may be that the size of the group has been decreased to such an extent as to decrease the size of the correlation coefficients. Selection, of course, also tends to reduce the size of the correlation.

TABLE XLIV.

Correlations Between Test and Examination Results for the Age Group

Characteristics Compared	Group Q		Group Q-A	
	r	N	r	N
	December			
I v Q	.832	1007	.834	758
I v Qe	.822	1016	.818	775
I v Qa	.701	1017	.706	772
E v Qe	.870	1006	.982	772
A v Qa	.811	1004	.840	767
	May			
I v Q	.866	1809	.863	1350
I v Qe	.856	1845	.852	1377
I v Qa	.745	1837	.744	1367
E v Qe	.861	1836	.859	1371
A v Qa	.808	1827	.838	1361

In/

In the case of the intercorrelations between test scores and examination marks there is very little difference between the results for the age group and the total qualifying group. Any increase which might have arisen seems to have been neutralized by the decrease in the size of the group.

TABLE XLV.

Correlations Between English and Arithmetic for the Age Group

Characteristics Compared	Group Q		Group Q-A	
	r	N	r	N
	December			
Qe v Qa	.685	1041	.687	793
E v A	.681	1023	.689	779
	May			
Qe v Qa	.742	1893	.750	1408
E v A	.724	1901	.735	1420

TABLE/

TABLE XLVI.

Correlations Between Teachers' Estimates and Intelligence Test and Qualifying Examination Results for the Age Group

Characteristics Compared	Group Q		Group Q-A	
	r	N	r	N
	December			
T v I	.697	1049	.705	795
T v Q	.786	1038	.797	787
	May			
T v I	.724	1941	.723	1450
T v Q	.767	1848	.770	1372

Tables XLV and XLVI show that there is no significant difference in the correlations when the age factor is removed by using a group in which the age is constant.

E. OBSERVATIONS

The general conclusion may be drawn that the inter-correlations between test scores and examination marks as given in Section III seem to be uninfluenced by the age factor.

On the whole, the results of this section are not in agreement with the results of other investigations. Thomson, for example, found that the correlations between age and each of his measures - intelligence test, scholastic test, qualifying examination, and teachers' estimates - were in the neighbourhood of/

of - ·35. The American investigators also obtained correlations with age to the extent of - ·36.

One important factor must be considered, namely, the compostion of the group on which the various correlations are based. The intercorrelations between the measures were based on two groups, the total group (T) and the Qualifying group (Q).

About 90% of the total group (T) lie within the age-group $11\frac{1}{2}$ - $12\frac{1}{2}$ years. Thus, when the correlation between age and a test score is found over group (T) it will be largely the correlation between these two variables within the age-group, $11\frac{1}{2}$ - $12\frac{1}{2}$ years. This small effective age-range may be a reason for the size of the correlations between age and test scores based on this group. Of course the 10%, because they are outsiders, may have an effect on the correlation which is out of proportion to their number.

By finding the intercorrelation of two test scores, say, $r_{I.E}$, over the age-group A, it was expected that the age factor would be eliminated. There was actually very little difference between $r_{I.E}$ over group (T) and the same correlation over Group A, ·891 as against ·892. It does not necessarily follow that the age factor is of little account as the reason for the similarity of the coefficients may be that there is little difference in the composition of Group T and Group A.

In the case of Group Q, 75% of the pupils lie within the age-group $11\frac{1}{2}$ - $12\frac{1}{2}$ years, and in addition it is a specially selected group. It may be, therefore, that the true correlation between/

between age and examination marks is masked by the operation of these two factors, the dominance of the age group and selection.

Removing the age factor from the intercorrelation between two measures, say $r_{I.Q}$, by finding the correlation over an age-group also presents difficulties. In the qualifying group Q the age-group has a dominating influence and so the similarity of the groups may be a reason for the correlation remaining unchanged. The diminution in the size of the group may also tend to neutralize the removal of the age factor. It is difficult to say, therefore, what influence the age factor actually possesses.

Probably the most significant correlations with age are those for the Qualifying pupils who are also in the age-group (Group Q-A). This group is homogeneous both in regard to age and in regard to scholastic opportunity. The fact that these correlations are low (about - ·1) seems to support the argument that the various measures correlate only to a slight extent with age.

V. THE RELIABILITY OF THE TESTS AND EXAMINATIONS

A. INTRODUCTORY

The reliability of a test or examination may be described as the consistency with which it yields its results. There are two aspects of reliability (1) the reliability of marking and (2) the reliability of the test. It is the second aspect with which we are concerned as the tests were objectively scored and therefore their reliability as regards marking is bound to be high. The Qualifying Examination was also, to some extent, objective in character.

In general use there are three methods [1] for

---

1. Statistics in Psychology and Education - H. E. Garrett Longman's Green & Co., 1937 - page 311

---

determining the reliability of a test

(1) Repetition of the test

(2) Use of parallel forms of the test

(3) Split-half method.

The first method consists simply of repeating the test and correlating the two sets of scores. When parallel forms of the test exist then the reliability is found by correlating the scores given by the two tests. In the split-half method the test is broken into two equivalent parts and, from the correlation of the half-tests, the reliability of the whole test may be calculated/

calculated. The second method is the one adopted in this investigation as parallel forms were purposely chosen for the December and May tests and examination.

As far as group intelligence tests and achievement tests are concerned a fairly high degree of reliability is expected.

"Most makers of general intelligence tests, and of educational achievement examinations, report correlations of ·90 or more between duplicate forms of their tests in unselected groups of fairly narrow age range." [1.]

---

1. Op. cit. page 315

Freeman may also be quoted as stating that "the reliability of mental tests is usually about ·70 to ·80". [2.]

---

2. Mental Tests, F. N. Freeman. Harrap & Co., Ltd.

Individual intelligence tests, also, have a high reliability, that of the latest revision of the Stanford-Binet Scale being given as ranging from ·85 to ·95 with a median of ·91. [3.]

---

3. Measuring Intelligence, Terman and Merill. Harrap & Co. 1937 page 47.

Examinations, particularly those of the essay type, have as a rule a much lower reliability. For example, in the School Certificate Examination in Mathematics the questions were divided into two parts (1) easy questions of a routine type (2)/

(2) questions of greater length and difficulty. An analysis of 236 scripts gave a value of r as ·255. Again, in a Junior County Scholarship for a large area for the Arithmetic test consisting of twenty questions the correlation between the odd and even was ·480 for 150 papers. [1].

---

1. The Necessary Imperfections of an Examination F. Sandon
Brit. J. Educ. Psych. Vol V, Part II, June 1935. page 180.

The question of the reliability of examinations has been fully dealt with by the English Committee of the International Examinations Enquiry Committee. The results of their investigations have been published in "An Examination of Examinations" and the "Marks of Examiners". In these publications they have shown how unreliable even carefully devised examinations with detailed instructions for marking may be.

B. RELIABILITY OF THE TESTS

(a) The Intelligence Tests

1391 of the pupils who were tested by M.H.T.20 in December, 1935, were also tested by M.H.T.21 in May, 1936. The correlation between the scores on these duplicate tests was

$$r = + \cdot 898 \pm \cdot 004 \quad (N = 1391).$$

As has already been shown the correlation between age and the scores in the tests is extremely low and thus the effect of 'partialling-out' age from this reliability coefficient is/

is practically negligible. When the age factor is removed the coefficient remains

$$r = .898.$$

(b) The English Tests

1348 of the pupils who were tested by M.H.E.7 in December, 1935, were also tested by M.H.E.8 in May, 1936. The correlation between the scores on the tests was

$$r = + .931 \pm .002 \quad (N = 1348).$$

The coefficient corrected for age was

$$r = .931.$$

(c) The Arithmetic Tests

1319 of the pupils who were tested by M.H.A.7 in December, 1935, were also tested by M.H.A.8 in May, 1936. The correlation between the scores was

$$r = .877 \pm .004 \quad (N = 1319).$$

This when corrected for age remained

$$r = .877.$$

These figures are in accordance with general expectation. The reliability of the English test, however, is remarkably high.

C. THE RELIABILITY OF THE QUALIFYING EXAMINATION

The examination marks of the representative sample of 500 pupils in Examination A and Examination B when correlated gave the reliability of the various parts of the Qualifying Examination.

(a)/

(a) The Total Qualifying Examination

The correlation between the total marks in the two examinations was

$$r = .933 \pm .004 \quad (N = 500),$$

with age partialled out

$$r = .932.$$

(b) The English Qualifying Examination

The correlation between the marks in English in the two examinations was

$$r = .920 \pm .005 \quad (N = 500),$$

with age partialled out

$$r = .918.$$

(c) The Arithmetic Qualifying Examination

The correlation between the marks in Arithmetic in the two examinations was

$$r = .864 \pm .008 \quad (N = 500),$$

with the age factor removed

$$r = .860.$$

The reliability of the Qualifying Examination is remarkable especially as the English Examination contained an essay paper.

In Examination A and Examination B the examiner classified the pupils into four categories (see Section I, page 120). Table XLVII shows the distribution of the 500 pupils in the two duplicate examinations.

TABLE/

TABLE XLVII.

The Distribution of Pupils in the Sample in Examination A and Examination B

Category	No. of Pupils in category in 1st examination	Distribution in Duplicate examination
A	126	A - 111 B - 14 C - 1 D - 0
B	82	A - 25 B - 36 C - 21 D - 0
C	157	A - 3 B - 22 C - 112 D - 20
D	135	A - 0 B - 0 C - 28 D - 107
Total	500	500

This table represents a correlation of ·933. Even with a correlation of this size the intermediate categories B and C show considerable divergences in the duplicate examinations.

D. OBSERVATIONS

The reliability of the Qualifying examination is remarkably high. This result must be considered from the two aspects of the reliability of an examination; the reliability of/

of the marking and the reliability of the examination itself.

One reason for the high reliability may be due to the fact that the examinations were corrected by a paid examiner who had a wide experience of the marking of Qualifying examinations throughout Scotland. Although no detailed scheme of marking was given, it seems likely that he would have a standard against which he assessed the value of each answer. Tribute must be paid to him for the care which he must have taken with the evaluation of the answers. Further experiments will have to be made so as to investigate the result of giving different examiners the same scripts to mark and of returning to the original examiner certain scripts to re-mark without any knowledge of his previous correction.

A reason for the high reliability of the examination itself may be due to the standardised conditions under which the examinations were set. Normally, this aspect of examinations is not stressed to the same extent as was done in the investigation.

The English examination was to some extent objective and most of the questions required only a short answer. On the other hand, it included an essay which is extremely difficult to assess with any degree of reliability. The field covered by the examination also was fairly wide including Spelling, Dictation, Composition, English Comprehension, Language Usage and the like, and thus divergences in any one section would not have any great effect/

effect on the total mark.

An Arithmetic examination of the type set is often very difficult to assess. The allocation of marks to an answer which is slightly wrong is extremely difficult, especially in the case of a problem. This is quite different from the test where one mark was given to each correct answer.

In making a comparison of the reliabilities of the tests and examinations some account must be taken of the duration of each. In Table XLVIII it can be seen that the English examination extended over a much longer period than any of the other tests or examinations.

TABLE XLVIII.

Duration of Tests and Examinations

Test	Duration
I	45 mins.
E	40 mins.
A	30 mins.
Qe	2 hours 20 mins. (app.)
Qa	1 hour 10 mins.

The chances are, therefore, that this examination if carefully devised will have a greater reliability than a test in the same subject lasting for less than a third of the time. The Arithmetic examination also lasts for a greater time than any of the tests.

Another/

Another point worth considering also is that the Dundee children were quite experienced in examinations of this type but had little or no experience with scholastic tests. It has been pointed out in "The Intelligence of Scottish Children" that experience with tests of intelligence does make a difference in the standard of achievement.

VI. THE PREDICTIVE VALUE OF THE INDIVIDUAL MEASURES

A.

The main purpose of this investigation was to find the measure or the combination of measures which would give the best prediction of success in the secondary school. It was agreed to measure success by the marks scored in the terminal examinations. Results were obtained from the four secondary schools controlled by the Dundee Education Committee, to which pupils from the various primary schools were promoted. These schools will be denoted by H, M, L and G respectively.

As has been pointed out there were two times for promotion to post-primary courses of instruction. Only one of these was operative in the case of the five year secondary schools which meant that the number of pupils in the first year of the secondary schools was composed partly from the December group and partly from the May Qualifying group. Actually the number of December qualifiers entering the secondary schools was comparatively small so these were omitted from the calculations.

The number of pupils in the first year group in each school is given in Table XLIX.

TABLE/

TABLE XLIX.

No. of Pupils in First Year Secondary Groups

School	No.
H	98
M	148
L	47
G	27
Combined Schools	320

Although the record card was the same for all schools each school was responsible for the organisation and administration of its own examinations. Each school, also, had its own method of determining the mark to be entered on the record card. In school H, for example, the marks were based solely on the terminal examinations while in school M a certain percentage of the marks in each subject was given for the class work during the term. The average mark for the term was an average of the marks for a number of specially selected subjects. These subjects varied from school to school although the main subjects of the course were generally common to all schools.

Within each school, too, the marks for the various subjects were not always comparable. In some subjects the one examination was set to all pupils and each class teacher corrected the complete paper. On the other hand, the possibility of different standards of marking was overcome in

some/

some cases by individual teachers being responsible for the correction of a single question in all the papers whether he had been teaching the classes or not. The problem was still further complicated by the fact that in some subjects different classes were given different examination papers.

Just as the subjects specially selected for the average mark varied from school to school, they varied also within a school according to the course the pupil was following. On the whole, however, the subjects common to all courses were numerous enough to predominate over one or two special subjects.

Despite the incomparability of the marks it was decided that in each school the average marks for each term would be taken as the measure of success in the course. Any standardisation or scaling of the marks would entail a great amount of labour and it is questionable whether such an undertaking would make a significant difference in the results.

B. CORRELATIONS BETWEEN THE VARIOUS MEASURES AND THE AVERAGE MARKS IN THE FIRST YEAR OF THE SECONDARY SCHOOLS

The correlations between the various measures and the average term examination marks were found for each of the three term examinations and for the average marks over the whole year. These correlations for school H are given in Table L.

TABLE/

TABLE L.

Correlation Between the Various Measures and First Year Marks in School H

Test	1st Term	2nd Term	3rd Term	Average for Year
I	.776	.678	.696	.737
I.Q.	.753	.669	.722	.728
E	.592	.526	.559	.583
A	.793	.704	.630	.679
Qe	.777	.757	.728	.772
Qa	.656	.649	.672	.673
T	.661	.664	.655	.682

From these results there is little difference in the predictive value of the various measures. The English test consistently gives the lowest correlation while the English examination on the whole gives the highest. On the other hand, the Arithmetic test has a superior predictive value to the Arithmetic examination. The results of the Intelligence test do not have a better predictive value than all the other individual measures as was the case with the majority of the other investigations. Teachers' estimates, despite the known variation in the standard from school to school, rank next to the results of the Intelligence test as far as the prediction of the marks over the whole year are concerned. This seems to support the conclusion that if these estimates could be scaled so/

so as to be on an objective standard, they would provide a useful means of estimating a child's future performance in the secondary schools' examinations.

Before a further analysis is made of these results it is essential to find out whether the tendencies shown in this school are also true of the other schools.

Table LI gives the results for school M.

TABLE LI.

Correlation Between the Various Measures and First Year Marks in School M

Test	1st Term	2nd Term	3rd Term	Average for Year
I	.715	.714	.680	.710
I.Q.	.696	.694	.703	.707
E	.580	.616	.586	.598
A	.700	.697	.698	.710
Qe	.743	.732	.731	.734
Qa	.662	.669	.695	.679
T	.716	.685	.679	.688

The coefficients in Table LI show a remarkable similarity to those in Table L. The superiority of the English examination over the English test, the superiority of the Arithmetic test over the Arithmetic examination and the relative value of the I.Q. and the teachers' estimates are evident.

The/

The results for the other two schools are given in Table LII and Table LIII.

TABLE LII.

Correlation Between the Various Measures and First Year Marks in School L

Test	1st Term	2nd Term	3rd Term	Average for Year
I	.578	.600	.676	.582
I.Q.	.546	.432	.487	.542
E	.432	.437	.648	.542
A	.489	.415	.533	.456
Qe	.668	.684	.774	.724
Qa	.604	.521	.653	.619
T	.393	.320	.363	.437

TABLE LIII.

Correlation Between the Various Measures and First Year Marks in School G

Test	1st Term	2nd Term	3rd Term	Average for Year
I	.769	.755	.750	.765
I.Q.	.673	.702	.691	.683
E	.699	.599	.645	.664
A	.781	.784	.764	.796
Qe	.781	.706	.758	.772
Qa	.755	.788	.759	.777
T	.846	.796	.831	.816

No/

No great reliance can be placed on the results for school L and school G since the number in the groups on which the correlations are based are small. A remarkable feature, for example, is that the teachers' estimates in school G give the highest coefficients to be found in these tables while for school L they give the lowest.

From these results it would appear that the best single measure for predicting the school marks of a pupil in the first year of a secondary school is the English examination. In this connection it should be noted that the examinations in the investigation were set under rigorously standardised conditions which makes the results much more comparable with those of the tests. Other investigations do not seem to have paid the same attention to this aspect of examinations. Several of the questions in the English examination were such as to allow of objective marking.

In comparing the various measures some account must be taken of the duration of the tests and examinations as was done with the question of reliability. The chances are that a carefully devised English examination lasting for approximately two hours twenty minutes will have a greater predictive value than a test in the same subject lasting for less than a third of the time.

A comparison of the questions in the examinations and tests also throws some light upon the results. The English examination/

examination with its papers in Spelling, Dictation, Composition and General English has the possibility of covering a wider field than the purely objective test. Composition, admittedly a difficult subject to assess, does measure an ability which cannot be measured to the same extent by the objective test. To some extent there will be a standardisation of the marking as the one examiner evaluated all the scripts, although Hartog showed that one examiner was by no means reliable in his evaluation.

The content of the Arithmetic tests and examinations shows clearly why the test is likely to be more efficient even though the latter are of longer duration. The examination consisted of ten mental problems and five questions, three of which were mechanical. On the other hand the test consisted of (1) over forty mechanical arithmetic questions some of which were similar to those in the examinations and (2) sixty problems ranging from simple questions as were given in the mental arithmetic paper to more complicated problems. Practically the whole of the content of the Arithmetic examination formed a small portion of the test.

The evaluation of questions in an Arithmetic examination raises another difficulty. To judge the seriousness of a mistake and to allocate marks accordingly is an extremely difficult problem. In the test, however, one mark is given to each correct answer and thus there is not the slightest dubiety about/

about the evaluation.

In view of these facts it is not surprising, therefore, that the Arithmetic test gives higher correlations than the Arithmetic examination.

There is not much difference in the results for the Intelligence test and the English examination. In the discussion on the intercorrelations between the various measures it was seen that these were measuring much the same ability. The Intelligence test depends to a considerable extent on reading comprehension.

The size of the correlations with the teachers' estimates are remarkably high in comparison with other investigations. It should be pointed out, however, that these estimates are practically a record of the child's work in the elementary school particularly in the Qualifying class. Many teachers summarized the marks in the examinations during the session and entered these on the appropriate schedules. The American investigators found that the elementary school record gave the best estimate of success in the first year of the junior high school and thus it might be concluded that, if these estimates could be adjusted to ensure that the marks were all on the same standard, they would provide a reliable means of predicting success in the first year of the secondary school.

The scaling of marks necessary to bring about a standardisation would entail, especially in large areas, a great deal/

deal of labour. Other more satisfactory means have been suggested. Among these are (1) the greater use of standardised tests in schools which would provide the teachers with an objective standard against which they could compare their own examinations and (2) the sending of the annual reports of the progress of the pupils in the secondary school to the primary schools in which they had been taught. This might assist the primary school teachers to understand to a greater extent what was necessary for success in the secondary school.

As would be expected there is a gradual decrease in the size of the correlations from term to term, although the decrease is slight. There is one exception, namely, the Qualifying Arithmetic examination. In every case there is an increase in the size of the correlation with the third term marks as compared with that given by the first term marks. It is difficult to say whether this is due to chance or not.

C. CORRELATION BETWEEN THE VARIOUS MEASURES AND THE AVERAGE MARKS IN THE FIRST YEAR : ALL SCHOOLS COMBINED

There are many objections to the combining of the results for the four schools in order to give a larger group on which to base the correlations. It has already been pointed out how many variable factors there are in the marks from school to school. On the other hand, there are several factors which tend to lessen the influence of these discrepancies. For example, the Leaving Certificate so dominates the secondary school/

school course that it gives a certain uniformity in all secondary schools; there is a tendency on the part of secondary school teachers to regard 50% as the pass mark in any subject. Chance factors will tend to neutralize each other when the marks are combined.

TABLE LIV.

Correlation Between Various Measures and First Year Marks in Combined Secondary Schools

Test	1st Term	2nd Term	3rd Term	Average for Year
I	.697	.646	.648	.683
I.Q.	.679	.626	.657	.666
E	.575	.556	.558	.577
A	.660	.640	.646	.666
Qe	.739	.714	.708	.729
Qa	.655	.636	.676	.672
T	.673	.626	.665	.680

The results in Table LIV bear out to some extent the conclusions which were drawn from the results for the individual schools. The English test gives the lowest correlation while the English examination gives the highest. There is little to choose between the predictive value of the Arithmetic test and examination though the former gives the higher correlations for the first two terms. As before, the Intelligence test gives correlations intermediate among those given by the other measures.

Over/

Over the year teachers' estimates rank second in predictive value to the English examination. An attempt was made to make these estimates more objective by the introduction of the Rating Card. It was found that there was little improvement in the predictive value of the estimates by this method. Table LV gives a comparison between the categoric and the numerical estimates.

TABLE LV.

Correlations Between Success in 1st Year Secondary School and Teachers' Estimates Given on a Categoric and Numerical Scale

	1st Term	2nd Term	3rd Term	Average for Year
5 point Scale	.541	.639	.638	.647
Numerical	.673	.626	.665	.680

D. CORRELATIONS BETWEEN THE VARIOUS MEASURES AND THE AVERAGE MARKS IN THE SECOND YEAR OF THE SECONDARY SCHOOLS

Results for the separate terminal examinations were not found, only those for the average mark over the year being calculated. The numbers in the groups were somewhat reduced as is shown in Table LVI.

TABLE LVI.

No. of Pupils in Second Year Secondary Groups

School	No.
H	78
M	129
L	39
G	24
Combined Schools	270

Table LVII gives the correlations between the various measures and the results of the second year examinations in school H. In order to give a comparison the corresponding first year results are also shown.

TABLE LVII.

Correlation Between the Various Measures and Second Year Marks for School H

Test	1st Year	2nd Year
I	.737	.615
I.Q.	.728	.691
E	.583	.381
A	.679	.579
Qe	.772	.626
Qa	.673	.631
T	.682	.533

In all cases the correlations for the second year, as might have been expected, show a decrease. This is least in the case of the I.Q. which for this school gives the best predictive value for the second year examination results. The English test, which gave the lowest correlation with the first year results, shows the greatest decrease among the various measures. It is surprising to note that the Arithmetic examination gives a higher correlation than not only the Arithmetic test but also the English examination. The teachers' estimates/

estimates show for the second year results a lower predictive value relative to the other measures than for the first year results.

The results for school M are given in Table LVIII.

TABLE LVIII.

Correlations Between the Various Measures and Second Year Marks for School M

Test	1st Year	2nd Year
I	.710	.653
I.Q.	.707	.599
E	.598	.581
A	.710	.675
Qe	.734	.636
Qa	.679	.665
T	.688	.632

The results for school M do not correspond entirely with those for school H. For example, the Intelligence test is only slightly superior to the English test as regards prediction. Perhaps the most striking feature of all is the small differences between the coefficients, the range being ·094. As in school H there is a general decrease in the size of the correlations. The Arithmetic examination though not superior to the Arithmetic test becomes superior in predictive value to the English examination. Teachers' estimates show the/

the same relative decline as in school H.

The results for schools L and G are given in Tables LIX and LX. No great importance can be attached to these results, however, as the numbers in the two groups are so small.

TABLE LIX.

Correlations Between the Various Measures and Second Year Marks in School L

Test	1st Year	2nd Year
I	.582	.712
I.Q.	.542	.761
E	.542	.561
A	.456	.501
Qe	.724	.721
Qa	.619	.552
T	.437	.489

TABLE LX.

Correlations Between the Various Measures and Second Year Marks in School G

Test	1st Year	2nd Year
I	.765	.734
I.Q.	.683	.675
E	.664	.638
A	.796	.791
Qe	.772	.679
Qa	.777	.801
T	.816	.797

The correlations between the various measures and the second year examination marks vary so much from school to school that it is difficult to draw any conclusion from them. A reason for these variations may be in the types of examination set in the different schools. 'Success' in one school may require a different ability for 'success' in another hence the measures would correlate with success to a different extent in different schools. For example, in school M the terminal marks in the different subjects entered on the record card are made up in some cases partly from terminal examinations and partly from a record of work throughout the term. This may give greater weight to a pupil's opportunity for home study than is the case in school H where the marks are mainly the results of the terminal examination. Such an explanation might account for the high correlation with I.Q. in school H and the relatively low correlation in school M.

Table LXI gives the results for the combined secondary schools.

TABLE/

TABLE LXI.

Correlations Between the Various Measures and Second Year Marks for the Combined Secondary Schools

Test	1st Year	2nd Year
I	.683	.603
I.Q.	.666	.563
E	.577	.516
A	.666	.615
Qe	.729	.622
Qa	.672	.672
T	.680	.590

In considering these results account must be taken of the fact that the number in the group for school M is practically half of the complete group. The tendencies, therefore, shown in school M will to some extent dominate in the larger group. Since, also, the coefficients vary to such an extent in their relative position to each other it seems that conclusions based on the results for the combined schools will require careful scrutiny.

The consistency in the size of the Arithmetic examination is striking and accounts for the highest correlation among the measures. The English test gives the lowest correlation in both years while the Intelligence test ranks next giving a slightly lower correlation than the teachers' estimates/

estimates. In all cases there is a decrease in the correlations with the second year marks as compared with those given with the first year marks.

E. GENERAL CONCLUSIONS

The correlations between the various measures and the secondary school marks in this enquiry are, on the whole, higher than those found in previous investigations. This may be due to some extent to the higher reliability of the tests and examinations used in this investigation, the more careful determination of these measures and the larger number in the groups on which the correlations are based. In the secondary school groups, also, there is likely to be a wider range of ability than in the groups used in the English investigations. Selection for entry into the English secondary schools is fierce whereas in the Dundee schools anyone, no matter the ability, could be admitted to the secondary departments if the necessary fees were paid.

The correlations are lower for the second year marks than for the first year marks. During the first year, however, there is very little fluctuation in the correlations for the various measures from term to term. This may be due to the consistency of the school marks. In school H, for example, the terminal examination marks when correlated gave coefficients well above .9.

The/

The range of the correlation coefficients for any term or year is small. Since the intercorrelations between the various measures are fairly high it may be considered that they are measuring the same abilities. If this were the case then the predictive value of the various measures would be similar.

The best individual measure for predicting school marks in the first year of the secondary school is the English examination. For the second year, the Arithmetic examination appears to give the highest correlations with secondary school marks. It may be deduced, therefore, that the Qualifying examination in English and Arithmetic will be a powerful agent in determining 'success' in the secondary school.

The English test has the poorest predictive value among the various measures both in the first and second years. In view of previous investigations the results of the Intelligence test rank surprisingly low among the measures and show no signs of increasing in predictive value for the second year.

Teachers' estimates take an intermediate position among the measures although their value might increase if some method of standardisation were adopted.

VII. THE PREDICTIVE VALUE OF COMBINATIONS OF MEASURES

The correlations between the various measures and 'success' in the secondary school were found so as to point to those which could be most usefully employed in predicting secondary school marks. Abundant evidence exists to show the danger of using a single measure as a means of selection or prediction. From the survey of previous investigations also it was seen that the highest coefficients were obtained by using a combination of measures.

In practice the individual measures in this investigation would rarely be used as the sole means of selection of pupils for post primary courses of instruction. For example, the English and Arithmetic tests or examinations are generally combined to form a single test or examination mark. Knowing the predictive value of the component parts of an examination, however, it is possible to point to the part on which improvement may be made.

The combinations discussed are those which would likely be chosen by an Education Committee in the selection of pupils for their post primary schools. Simple combinations such as Qe + Qa and E + A immediately suggested themselves. To these were added separately and then together I.Q. and T.

The results for school H are given in Table LXII.

TABLE/

TABLE LXII.

Correlation Between Combinations of Measures and First Year Marks in School H

Tests	1st Term	2nd Term	3rd Term	1st Year Average
Qe + Qa	.794	.779	.772	.801
E + A	.768	.682	.659	.700
I.Q. + Qe + Qa	.828	.786	.801	.823
I.Q. + E + A	.812	.721	.725	.756
Qe + Qa + T	.795	.786	.778	.808
E + A + T	.798	.739	.719	.758
I.Q. + Qe + Qa + T	.817	.786	.795	.819
I.Q. + E + A + T	.824	.755	.767	.787

From Table LXII it is evident that the combination of the English and Arithmetic examinations gives a more accurate prediction than a similar combination of tests. When the I.Q. is added to the Qualifying examination mark (Qe + Qa) the combined mark gives the most accurate prediction of all. The effect of adding the teachers' estimates to this mark, for example, is to lower the correlation coefficient.

It is difficult to tell whether the I.Q. or the teachers' estimates combines better with the English and Arithmetic combined test score. Over the whole year the teachers' estimates give a correlation which is ·002 higher than that given by the I.Q. in combination with the tests. By combining all four, I.Q., E, A and T, a mark is obtained which is better than any of those which/

which involve E and A though not so good as any which have Qe and Qa in the combination.

Actually there is not a great difference between any of the coefficients. For the first term examinations the range is ·060 while for the average marks over the whole year the range increases to ·123.

The corresponding results for school M are given in Table LXIII.

TABLE LXIII.

Correlation Between Combinations of Measures and First Year Marks in School M

Tests	1st Term	2nd Term	3rd Term	1st Year Average
Qe + Qa	.773	.771	.785	.778
E + A	.707	.725	.709	.723
I.Q.+ Qe + Qa	.784	.781	.794	.796
I.Q.+ E + A	.731	.743	.735	.745
Qe + Qa + T	.781	.775	.775	.773
E + A.+ T	.766	.767	.753	.766
I.Q.+ Qe + Qa + T	.790	.779	.787	.787
I.Q.+ E + A + T	.770	.770	.763	.773

The results in Table LXIII show exactly the same tendencies as those in Table LXII. The Qualifying examination mark has a better predictive value than the combined English and Arithmetic test score. I.Q's give a higher correlation with the Qualifying/

Qualifying mark than do the teachers' estimates but the reverse is the case with the combined test score. Teachers' estimates raise the correlations given by I.Q. + E + A but lower those given by I.Q. + Qe + Qa.

A remarkable feature about the coefficients in this table is their similarity. The lowest coefficient in the table is ·707 while the highest is ·796. In magnitude they agree to a large extent with those given in Table LXII.

Tables LXIV and LXV give the corresponding results for schools L and G.

TABLE LXIV.

Correlations Between Combinations of Measures and First Year Marks in School L

Tests	1st Term	2nd Term	3rd Term	1st Year Average
Qe + Qa	.757	.717	.849	.799
E + A	.586	.542	.752	.636
I.Q. + Qe + Qa	.789	.711	.831	.819
I.Q. + E + A	.640	.560	.728	.672
Qe + Qa + T	.696	.638	.748	.744
E + A + T	.617	.551	.725	.674
I.Q. + Qe + Qa + T	.733	.649	.754	.769
I.Q. + E + A + T	.650	.561	.710	.691

TABLE/

TABLE LXV.

Correlations Between Combinations of Measures and First Year Marks in School G

Tests	1st Term	2nd Term	3rd Term	1st Year Average
Qe + Qa	.869	.845	.858	.876
E + A	.838	.783	.798	.827
I.Q.+ Qe + Qa	.860	.855	.860	.869
I.Q.+ E + A	.816	.790	.796	.812
Qe + Qa + T	.895	.860	.882	.888
E + A + T	.893	.836	.859	.873
I.Q.+ Qe + Qa + T	.893	.875	.889	.891
I.Q.+ E + A + T	.872	.838	.852	.860

The results in Table LXIV confirm the previous conclusions even though the number in the group is fairly small. In school G where the group numbers only 27 and hence makes the results unreliable the best measure is given by I.Q.+ Qe + Qa + T.

TABLE/

TABLE LXVI.

Correlations Between Combinations of Measures and First Year Marks for the Combined Secondary Schools

Tests	1st Term	2nd Term	3rd Term	1st Year Average
Qe + Qa	.784	.759	.778	.788
E + A	.697	.675	.679	.702
I.Q. + Qe + Qa	.790	.754	.779	.788
I.Q. + E + A	.728	.693	.707	.726
Qe + Qa + T	.788	.754	.782	.794
E + A + T	.755	.720	.740	.760
I.Q. + Qe + Qa + T	.798	.756	.786	.798
I.Q. + E + A + T	.766	.725	.748	.766

The results for the combined secondary schools in Table LXVI vary slightly from those of the individual schools. As regards the combined English and Arithmetic test score the position is the same. Teachers' estimates when added to the combined test score increase the correlation to a greater extent than the addition of the I.Q.. When I.Q., E, A and T are combined the average correlation with the first year marks is about ·77 which is greater than a combination of any of these four measures.

The Qualifying examination marks give a higher correlation with the respective first year marks than any of the combinations involving the test scores. The addition of either the/

the I.Q. or the teachers' estimates makes very little difference to the size of the correlation. When both are added this combined mark generally gives a slightly higher correlation than the Qualifying examination mark alone. The average correlation of this combination I.Q. + Qe + Qa + T is about ·79.

Correlations between these combinations of measures and the average mark for the second year for the individual schools are given in Table LXVII.

TABLE LXVII.

Correlations Between Combinations of Measures and Second Year Marks in the Various Schools

Tests	School H	School M	School L	School G
Qe + Qa	.696	.716	.760	.837
E + A	.532	.694	.676	.809
I.Q.+ Qe + Qa	.738	.709	.884	.839
I.Q.+ E + A	.627	.686	.796	.797
Qe + Qa + T	.679	.748	.737	.855
E + A + T	.583	.724	.729	.854
I.Q.+ Qe + Qa + T	.712	.710	.836	.863
I.Q.+ E + A + T	.643	.711	.808	.844

The correlations with the second year results show a certain decrease although this is not as great as might have been expected. In all the results with the second year examinations the factor of selection is at work. The majority of those who leave/

leave before completing the second year are pupils who have found the course too stiff. Pupils who are retarded account for some of the leakage also and thus accentuate the same selective factor. The content of the secondary schools examinations becomes more specialised and more likely to demand the use of abilities which are not measured by tests or examinations in English and Arithmetic. Reasons such as these could be put forward as explanations of a decrease in the size of the coefficients.

The superiority of the Qualifying examination is still evident but the variations from school to school make it difficult to draw any conclusion as to which combination gives the best prediction. Actually, in most cases, the addition of either or both I.Q. and teachers' estimates makes no great difference to the correlation particularly in the case of the Qualifying examination.

This conclusion is borne out when the results of the combined secondary schools are considered.

TABLE/

TABLE LXVIII.

Correlations Between Combinations of Measures and Second Year Marks for the Combined Secondary Schools

Test	r
Qe + Qa	.727
E + A	.639
I.Q.+ Qe + Qa	.708
I.Q.+ E + A	.644
Qe + Qa + T	.719
E + A + T	.680
I.Q.+ Qe + Qa + T	.711
I.Q.+ E + A + T	.676

In Table LXVIII the range of the coefficients is ·088 which shows that there is no great difference in the predictive value of the various combinations. The highest correlation is given by the Qualifying examination mark alone. When either or both the I.Q. and teachers' estimates are added to this mark there is a decrease in the size of the correlation. These measures increase the predictive value of the combined test score, the I.Q. to a less extent than the teachers' estimates.

The following conclusions may be drawn from the results in this section.

The predictive value of the various combinations of measures is greater than that of the individual measures. An even/

even better prediction would have been obtained if the inter-correlations between the tests and examinations had not been so high. In all cases of selection the dangers of using a single measure are great, particularly at the borderline, hence some combination of different measures is advisable in selecting pupils for a secondary school course.

The Qualifying examination used in this investigation has a better predictive value than the combined English and Arithmetic test results. This is due mainly to the high predictive value of the English examination although during the second year the Arithmetic examination becomes increasingly important.

For the first year, the combined results of the Qualifying examination, intelligence test and teachers' estimates give the best prediction of 'success' in the secondary school.

In the second year the Intelligence test and the teachers' estimates when added to the Qualifying examination mark do not improve the predictive value of that mark.

VIII. THE PREDICTIVE VALUE OF WEIGHTED COMBINATIONS OF MEASURES.

In the previous section it was seen that the various measures could be combined to give a higher correlation with secondary school marks than any individual measure. Each component of the combination has a predictive value of its own and thus it might, in some cases, be advisable to give greater weight to some measures than to others. For example, in the first year the English examination correlates more highly with the secondary school marks than the Arithmetic examination. It might, therefore, be advantageous to give greater weight to the English than to the Arithmetic in the Qualifying examination mark when predicting the first year secondary school marks.

The weight of an examination mark has no necessary relation to the total 'possible' mark in that examination. If the total possible mark in English is 150 and that in Arithmetic is 100 it does not follow that the respective weighting of these marks when combined is 3 : 2. The weight of a mark is controlled by the standard deviation of the marks in that subject.

To find the best weighting of measures in a combination recourse must be made to the regression equation. The method used to determine the weighting was one devised by Professor Thomson/

Thomson[1] and described in an article in the British Journal

1. On the Computation of Regression Equations, Partial Correlations, etc.. G. H. Thomson. Brit. J. of Educ. Psych., Vol. 23, Part 1, 1932.

of Psychology.

Only the correlations for the combined secondary schools with the average marks for the first and second years were found.

TABLE LXIX.

Correlations Between the "Best weighted" Combination of Measures and First Year Marks for the Combined Secondary Schools.

Tests	r(weighted)	weighting	un r(weighted)
Qe + Qa	.790	3 : 2	.788
E + A	.709	3 : 5	.702
IQ + Qe + Qa	.797	1: 3 : 2	.788
IQ + E + A	.733	5: 3 : 7	.726
Qe + Qa + T	.801	5: 3 : 2	.794
E + A + T	.766	2: 2 : 3	.760
IQ + Qe + Qa + T	.807	2: 6 : 4 :3	.798
IQ + E + A + T	.774	3: 3 : 4 :6	.766

In the above table the correlations for the corresponding unweighted combination of tests are given for purposes of comparison.

Perhaps/

Perhaps the most striking feature of the table is the small differences between the weighted and corresponding 'unweighted' coefficients. In every case the difference is in the third decimal place. This may be due to the fact that various measures correlate fairly equally with the criterion and highly with each other.

The highest coefficient is obtained by the combination IQ + Qe + Qa + T weighted in the proportions 2 : 6 : 4 : 3. This is only an increase of ·009 on the 'unweighted' coefficients. Contrary to conclusions of previous investigations, the I.Q. has the least weight in this combination.

By weighting the Qualifying examination mark by 3 : 2 in favour of the English examination the fairly high correlation of ·790 is obtained. The addition of the I.Q. and the teachers' estimates brings only an improvement of ·017. A similar addition tothe combined test score brings an improvement of ·065.

The relative placings of the various combinations as regards prediction of first year secondary school marks is exactly the same as for the unweighted combinations. The range of the coefficients is also small being ·098.

TABLE/

TABLE LXX.

Correlations Between the "Best Weighted" Combination of Measures and Second Year Marks for the Combined Secondary Schools.

Tests	r(weighted)	weighting	r(unweighted)
Qe + Qa	.730	2 : 3	.727
E + A	.648	1 : 2	.639
IQ + Qe + Qa	.729	1 : 15 : 21	.708
IQ + E + A	.653	1 : 1 : 2	.644
Qe + Qa + T	.731	5 : 6 : 1	.719
E + A + T	.683	3 : 4 : 4	.680
IQ + Qe + Qa + T	.730	1 : 5 : 7 : 1	.711
IQ + E + A + T	.683	1 : 3 : 4 : 4	.676

The results for the second year, given in Table LXX, show the same agreement between the weighted and the unweighted coefficients. There is, also, only a slight decrease in the size of the coefficients for the two years, although the weighting in the various combinations has changed.

Practically no alteration is made in the predictive value of the Qualifying examination mark by combining with it the results of the intelligence test or the teachers' estimates. The weight given to the intelligence test result when combined with the qualifying examination mark is extremely small.

A/

A striking change has taken place in the relative weighting of the English and Arithmetic examinations. In the first year the English examination received the greater weight while in the second year the reverse is the case.

From these results, the relatively high predictive value of the Qualifying examination mark is emphasised. It is also evident that any system of weighting is of doubtful practical value. In the selection of pupils for secondary schools an unweighted combination of measures would give almost as accurate results as a weighted combination and, from the point of view of an Education Committee, the adopting of the former would mean an appreciable saving in time and labour.

IX. THE RELATIVE PREDICTIVE VALUE OF THE QUALIFYING EXAMINATION and INTELLIGENCE TEST.

In the survey of previous investigations it was shown that, in the majority of the experiments, an intelligence test gave a more accurate prediction of post-primary school success than tests or examinations in English and Arithmetic. The position has been summed up in the following manner. "The prognostic value of the intelligence test is higher than that of the ordinary written test in either Arithmetic or English, but is sometimes very little higher."[1.]

---

1. The Selection of Children for Secondary Education. Davies & Jones. Harrap. 1936. Chap. 3.

The table in Part I, Section III, giving a summary of the results of the English investigations shows that the average correlation between secondary school marks and I.Q. was ·44 while the corresponding figure for the examination in English and Arithmetic was ·32.

A comparison between the predictive value of the intelligence test and qualifying examination for school H is given in Table LXXI.

TABLE/

TABLE LXXI.

Predictive Value of I.Q. and Q. for School H.

Test	Correlation Coefficients				
	1st Term	2nd Term	3rd Term	1st Year Average.	2nd Year Average.
Q.	.794	.779	.772	.801	.696
I.Q.	.753	.669	.722	.728	.691

From Table LXXI it is evident that the Qualifying examination used in this investigation has a better predictive value than the intelligence test as far as the first two years of a five year secondary course is concerned. The difference between the two measures, however, is slight.

The coefficients are high compared with the results of previous investigations. As this point has already been considered in connection with the correlations given by the individual measures, no further discussion will be made in this section.

There is little difference between the results from term to term over the first year. At the end of the second year, however, although the coefficients for both I.Q. and Q. are lower than for the first year, the fall in the case of the former is much smaller. The two coefficients are practically identical.

TABLE LXXII.

Predictive Value of I.Q. and Q. for School M.

Test	Correlation Coefficients				
	1st Term	2nd Term	3rd Term	1st Year Average.	2nd Year Average.
Q	.773	.771	.785	.778	.716
I.Q.	.696	.694	.703	.707	.599

The results for school M, given in Table LXXII, confirm the superior predictive value of the Qualifying examination. There is little variation, also, in the results over the first year but the fall in magnitude of the coefficients at the end of the second year is much more marked in the case of the I.Q. than for the Qualifying examination mark.

TABLE LXXIII.

Predictive Value of I.Q. and Q. for School L.

Test	Correlation Coefficients				
	1st Term	2nd Term	3rd Term	1st Year Average.	2nd Year Average.
Q	.757	.717	.849	.799	.760
I.Q.	.546	.432	.487	.542	.761

TABLE LXXIV.

Predictive Value of I.Q. and Q. for School G.

Test	Correlation Coefficients				
	1st Term	2nd Term	3rd Term	1st Year Average.	2nd Year Average
Q	.869	.845	.858	.876	.837
I.Q.	.673	.702	.691	.683	.675

Results for Schools L and G, given in Table LXXIII and Table LXXIV respectively, although no great reliance can be placed upon them, show that Q gives higher correlations with secondary school marks than does the I.Q. Table LXXV gives the results for the combined secondary schools.

TABLE LXXV.

Predictive Value of I.Q. and Q. for the Combined Secondary Schools

Test	Correlation Coefficients				
	1st Term	2nd Term	3rd Term	1st Year Average.	2nd Year Average.
Q.	.784	.759	.778	.788	.727
I.Q.	.679	.626	.657	.666	.563

The results in Table LXXV agree largely with those for school M. At the end of the second year, the difference between I.Q. and Q. is greater than that found during the first year.

From these results it seems clear that the Qualifying examination gives a better prediction of secondary school success than the intelligence test for the first and second year of the course. The evidence as to whether the difference in the predictive value of the Qualifying examination and intelligence test increases or decreases during the second year is most conflicting and no definite conclusion can be drawn.

The type of Qualifying examination used in this investigation may, to some extent, explain a conclusion which differs/

differs from that reached in the majority of the previous experiments. Many of the questions were of the objective type and others were framed in such a way as to admit of brief answers. The correction was undertaken by a paid examiner who corrected all the scripts and thus there was a greater chance of standardisation. This is evident from the reliability of the examination which is actually higher than any of the standardised tests, an unusual result in itself.

Another point on which too great stress cannot be laid is the fact that the examination was administered under a rigidly standardised procedure similar to the intelligence test. These variables were seldom controlled to such an extent in any of the reported investigations.

It was shown that the intelligence test and the English Qualifying examination may measure to a large extent the same ability, the correlation between the two being ·839. The Arithmetic examination, also, correlates fairly highly with the intelligence test, r being ·723. As the relationship between the English and Arithmetic examinations is not extremely high, these two may be taken as measuring different abilities which are measured by the intelligence test. By combining the two examinations, therefore, it might be expected that they would measure much the same abilities as the intelligence test and/

and in addition, abilities which are necessary for success in secondary school examinations.

Referring back to a previous investigation[1], two

---

1. Intelligence Tests and Entrance Examinations as Instruments for Selecting and Grading Students. C. Jorgenson.

---

important conclusions were drawn (1) that in the prediction of success in the post-primary school "the superiority of the intelligence tests is more marked among pupils of lower ability". (2) The main thing to ask of any entrance test is that it shall exclude those who would be least successful if admitted to a school. In one respect the entrance examination does seem to have a definite superiority: it appears to pick out the outstanding pupils better than does the intelligence test."

In order to test the validity of these conclusions the follow-up correlations of the different sections in the two large secondary schools were calculated. Pupils in school M were placed in A section if thought capable of studying two languages; otherwise they studied one language in Section C. In school H class IA was for those taking a two language course (French and Latin) and generally enrolled the brightest pupils in the school; Class IE was for pupils whose primary school record warranted only a one language course.

The/

The following table shows the rank correlations for school M.

TABLE LXXVI.

Rank Correlations with Success in the First Year of a 5 - year Secondary Course for A and C sections separately for School M.

Correlations between Success and	1st Term		2nd Term		3rd Term		Average for Year	
	IA	IC	IA	IC	IA	IC	IA	IC
Q	.73	.64	.73	.57	.68	.54	.64	.56
I	.51	.70	.51	.63	.49	.61	.58	.67

From the above table there seems little doubt but that the qualifying examination picks out the better pupils, while the intelligence test picks out the less able. This conclusion is confirmed by the results for school H.

TABLE LXXVII.

Correlations between Success and	1st Term		2nd Term		3rd Term		Average for year	
	IA	IC	IA	IC	IA	IC	IA	IC
Q	.78	.36	.77	.50	.83	.46	.83	.52
I	.84	.73	.72	.75	.80	.61	.81	.74

In the above table the superiority of the Qualifying examination over the Intelligence Test for the A class is less marked than in Table LXXVII.

From/

From the above results the conclusion may be drawn that, in this experiment, the Qualifying Examination has a better predictive value than the Intelligence Test. On the other hand, the Qualifying Examination selects the clever pupils better than does the Intelligence Test while the latter picks out the less able children more efficiently than the Qualifying Examination.

X. PREDICTIVE VALUE OF FIRST YEAR TERMINAL EXAMINATIONS IN THE SECONDARY SCHOOL

In several of the previous investigations, it was shown that the marks at the end of the first year gave a fairly reliable index of a pupil's likelihood of success at a later stage in a post-primary course. The correlations between marks gained in the first year and those at a later stage of the secondary school course were higher than those given by the entrance tests to that course. From these results it might be inferred that pupils should be given a probationary period in the first year of the secondary school.

The probationary period suggested varies from a few days to a complete year. In the Schulerauslese it is suggested that pupils from the fundamental school should be given a probationary period of eight to ten days before sitting the entrance test for that school. Dr. Earle states that a satisfactory prediction may be made after one term. The majority of the investigators, however, give one year as the probationary period.

Valentine supports the argument for a one year probationary period, the correlation between first year marks and those three years later being ·67. The Swedish investigators gave correlations for similar data as ·92 for five years later and ·93 for four years later. Collman results were ·88 in a three year period and ·86 in a five year period. American research/

research gave similar conclusions, the average correlation being approximately ·69.

First Year Marks as a Means of Predicting Future Success in the Secondary School

The following table gives the correlations between the average mark for the first year and the average mark for the second year in the various secondary schools.

TABLE LXXVIII.

Correlation Between First and Second Year Average Marks

School	r	N
H	.910	78
M	.878	129
L	.932	39
G	.829	24
Combined Schools	.893	270

These coefficients are higher than those given by any of the tests or combination of tests. It may be concluded, therefore, that by the end of the first year a fairly reliable prediction can be made as to a pupil's success in the next year of the course. In addition, this prediction will be more accurate than could have been made on the basis of the results of any of the tests and examinations given before entrance to the secondary school.

To find whether the probationary period of one term would be sufficient the correlations between the first term marks and the second year average marks were calculated. These are given in Table LXXIX.

TABLE LXXIX.

Correlation Between First Term and Second Year Average Marks

School	r	N
H	.864	79
M	.844	127
L	.904	38
G	.737	23
Combined Schools	.838	267

These correlations, on the whole, are not so high as those for the first year average marks but are higher than those for the tests and examinations.

There is definite increase in the correlations between the second year average marks and the second term marks as compared with the first term marks.

TABLE/

TABLE LXXX.

Correlation Between Second Term and Second Year Average Marks

School	r	N
H	.919	76
M	.883	127
L	.889	34
G	.713	25
Combined Schools	.878	262

Table LXXX gives the correlations between the second term and second year average marks.

Both the first term marks and the first year average marks correlate with the second year average marks to a greater extent than any combination of tests and examinations. This supports the conclusion that the most accurate means of predicting success in the secondary school is the first year term examinations of the secondary school. As has been pointed out, this implies a probationary period.

Since the correlations with the first year average marks are higher than the corresponding results with the first term marks it may be concluded that a more reliable prediction of later success in the course is given at the end of the first year than at the end of one term. This is further supported by the fact that the coefficients given by the second term marks approach closely to those given by the first year average marks. Such a result is not surprising as some children take a term or so to settle/

settle down in the new surroundings and organisation of a secondary school. There are others, particularly those who have been in the primary department of the same school, who have done fairly well in the primary school and who achieve a certain amount of success in the relatively simple first term examinations but who gradually lose ground as soon as the work becomes more difficult.

The high correlations given in Tables LXXVIII, LXXIX and LXXX may to some extent be due to a 'halo' effect. Clever pupils are unconsciously marked high while the less able children are often not given sufficient credit for their efforts. Despite the fact that the teachers normally vary from subject to subject and from year to year there is a prestige attached to a good pupil which stands him in good stead throughout his school course.

There are some reasons for expecting the first year secondary school marks to give higher correlations than any of the tests or examinations. It has been questioned, for example, whether tests or examinations in English and Arithmetic given in the primary school are capable of predicting the attainments of a child in a secondary school course which involves other special abilities.

Success in the secondary school examinations is often influenced by factors which do not have the same weight in the primary school. Children who receive special coaching and have the advantage of studying under very favourable conditions often do/

do better than their intelligence warrants. Usually in the later stages of the secondary school course homework and individual effort plays a greater part.

Furthermore the first year secondary school marks have the advantage of being determined under similar conditions to the later marks of the course. The buildings, the system of examination and the teaching are all constant factors. In the case of the tests and examination results, these have been determined under different conditions according to the primary schools from which the various secondary groups were selected.

The real problem, of course, is to find out the measure which has the best predictive value not for the first year but for the third or fifth year of the course. Naturally conclusions on this question cannot be made at this stage of the investigation. It may be that new factors will arise which will play a greater part in determining success in the secondary school course. On the other hand, other factors, such as the I.Q., may have a greater weight in the performance of a pupil at a later stage in the course. The likelihood is that, for the three years course at least, the marks in the first year of the course give the best prediction for success in the secondary school. This implies that a probationary period is advisable.

XI. THE BURSARY EXAMINATION

I. It has been shown that several investigators have proved the lack of validity of a single examination for the selection of children for secondary schools. Valentine[1], for example, showed that pupils who were just above the pass mark of an entrance examination often did remarkably well in their subsequent secondary school course. Conversely, pupils who did well in the entrance examination often failed to justify their entry into the secondary school. From these facts it is a safe deduction that some pupils who just failed to reach the pass mark might have made a success of a secondary school course. Naturally this is difficult to prove as these pupils are not given an opportunity of enrolling in a secondary school.

1. Reliability of Examinations.

Before an examination can have value as an instrument for prognosis it must be reliable, both in itself and in the marking. Generally, this is only possible if the examination is of the objective type. Hartog, for example, showed that even with detailed correction instructions the evaluation of an examination of the essay type is subject to variation.

II. These examinations are of the type which on the whole do not lend themselves to objective marking. Unfortunately there was no method by which the reliability of these examinations could be found.

The possible marks for the examinations are given in Table LXXXI.

TABLE LXXXI.

Possible Marks in Bursary Examinations.

	English	Arithmetic	Total
December	63	60	123
April	80	70	150

Pupils who scored 65% of the total possible marks in the examination were awarded bursaries. This method of selection naturally lays itself open to criticism. A mark or so makes all the difference between a 'pass' or a 'fail' and it has been clearly shown that such small variations in marks are in no wise a measure of such a big distinction of ability.

The following table gives the number of candidates gaining bursaries.

TABLE LXXXII.

Number of Bursary Candidates gaining Awards.

Bursary Examination	No. of candidates	No. of awards	% of awards
December	141	53	31
April	286	136	48

The percentage of candidates gaining bursaries was less in December than in April, which may mean that it was more difficult to gain an award on the one examination than on the other. If the standard of attainment necessary for an award was different for the two examinations then another factor causing discrepancies is introduced.

TABLE LXXXIII.

Mean and Standard Deviation of Bursary Examination Marks.

Examination	December		April	
	Mean	σ	Mean	σ
English	37	8.5	49	8.6
Arithmetic	36	12.7	44	13.2
Total	75	16.7	90	18.2

To gain a bursary on the April examination a pupil had to score a mark of $97\frac{1}{2}$ which was $7\frac{1}{2}$ marks above the mean of the marks made by the candidates. In the case of the December examination the pass mark was 80 which was 5 marks above the mean mark. Thus it would appear as if a candidate had an easier chance of gaining an award in December than in April. This conclusion, of course, implies that the two groups sitting the examination were of the same ability. Actually the April group was superior to the December group as regards intelligence the mean I.Q.s being 116 and 113 respectively.

From/

From the standard deviations of the marks, also, it may be concluded that the difference of a mark in the December examination is more significant than the corresponding difference in the April examination. The pass marks 97½ and 80 when converted into standard scores become ·41 and ·35 respectively which still seems to indicate slightly more difficulty in gaining a bursary in April than in December.

It is interesting to note that four candidates in December and ten in April scored exactly 65% of the marks. Eight candidates in December and nine in April were within one mark of the percentage.

The bursary examination candidates formed a small selected group of the total Qualifying Classes. Pupils whose chances of success were remote were usually discouraged from attempting the examination; a number of clever children did not wish secondary education; many children were prevented from applying for a bursary owing to their parent's income being above the level laid down in the regulations. Any correlation, therefore, involving the bursary examination results is likely to be affected by these selective factors. This is illustrated by comparing the correlations between the Qualifying examination marks and the scholastic test scores for the total Qualifying group and for the bursary group. The respective figures are ·91(973) as against ·75(141) for December

and/

and ·89(1764) as against ·72(286) for April. The probable error of the correlations for the bursary group is larger, also, than that for the total Qualifying group.

The correlations between bursary examination marks and Qualifying examination marks are ·59 for December and ·54 for April. These coefficients are relatively high considering the select nature of the group on which the correlations are based. It has been pointed out, however, that such correlations mean considerable discrepancies between the results.

Teachers' estimates agree to no better an extent with the bursary examination results, the correlations being ·45 for December and ·55 for April. In this case,not only is the selective factor having an effect but also the fact that the estimates are given by different teachers from different schools in which the standards are not likely to be the same.

The correlation of the bursary examination results with I.Q. gives a similar result, the respective correlations being ·60 for December and ·50 for April.

From these figures it seems evident that the results of a single examination of the type given in this connection do not agree with other assessments of the pupils. If the December Qualifying examination had been used as the means of selection then 19 of the pupils who failed to gain one of the 53 awards would have been successful. In the case of the May examination

32 failures would have been placed in the first 136 and would have been awarded bursaries.

Similarly 20 in December and 31 in April who were not awarded bursaries would have been successful if the awards had been granted on the basis of teachers' estimates. The numbers of failures gaining awards if the I.Q. had been taken as the criterion were 22 for December and 29 for April.

For the bursary candidates there were, in addition to the bursary examination results, intelligence test results, scholastic test results, qualifying examination marks and teachers' estimates. The bursary awards might have been made on five different marks, namely, I.Q., E + A, I + E + A, Q, and T. It was found that 9 of the 53 successful candidates in December and 6 out of 136 in April would have been accounted failures on <u>all</u> these five criteria.

Such results emphasise the lack of agreement between the bursary examination and the other methods of assessment. Only a follow-up, however, can say which of the measures has the greatest validity.

It is more difficult to prove that a pupil who had failed would have been successful than to prove that a pupil who gained a bursary justified his selection. There were, however, one or two pupils who failed to gain a bursary but proceeded to a secondary school at their own expense. Table

LXXXIV gives the follow-up data for three such cases.

TABLE LXXXIV.

Secondary School Record of Pupils who Failed in the Bursary Examination.

Pupil	A	B	C
Rank in Bursary Examination.	68.5 (Dec.)	152 (May)	222 (May)
I.Q.	127	121	124
Q.	211	172	185
E + A	202	163	164
T	92	90	*-
Secondary School Marks			
I(i)	85	79	76
I(ii)	85	79	78
I(iii)	83	79	75
I Av.	84	79	76
II Av.	84	77	70

* No return for teachers' estimates.

‡ Average for two terms. Pupil absent one term.

These cases illustrate how a single examination of this type may prevent pupils of even outstanding ability from gaining entrance to a secondary school. In Table LXXXIV, all the other test and examination results are in agreement with the secondary school/

school performance of the pupils.

Although it does not necessarily follow, it is probable that pupils who do well in the Central school would have successfully undertaken a secondary course. Several pupils who failed in the bursary examination but whose record on the other tests would have justified selection, made excellent progress in the Central schools to which they had been promoted. The likelihood is that they would have justified the award of a bursary. Table LXXXV gives examples of some pupils of this type.

TABLE LXXXV.

Central School Record of Pupils who Failed in the Bursary Examination

Pupil	D	E	F	G	H	I
Rank in Bursary Examination	64 (Dec)	80 (Dec)	152 (April)	137 (April)	146 (April)	159 (April)
I.Q.	121	125	129	138	129	132
Q.	227	203	211	185	160	201
E + A	189	181	188	182	164	190
T	89	83	83	83	95	90
Central School Marks.						
I(i)	80	82	82	81	81	77
I(ii)	76	83	85	82	76	75
I(iii)	78	85	82	78	80	75
I (Av)	78	83	83	80	79	75
II(Av)	75	84	79	85	82	77

Pupils/

Pupils who gained bursaries but whose selection is not justified according to the other methods of assessment, generally failed to make a success of the secondary school course. Table LXXXVI gives examples of pupils of this type and their secondary school record.

TABLE LXXXVI.

Secondary School Record of Pupils who Failed to Justify their Awards.

Pupil	K	L	M	N	O
Rank in Bursary Examination	39 (Dec)	43 (Dec)	51 (Dec)	82 (April)	123.5 (April)
I.Q.	103	101	102	101	101
Q.	145	141	171	161	157
E + A	159	108	152	121	144
T.	72	82	79	82	80
Secondary School Marks					
I(i)	44	60	59	64	53
I(ii)	49	57	54	59	54
I(iii)	57	55	61	58	57
I(Av)	50	57	59	60	54
II(Av)	60	50	53	48	50

In all these cases the I.Q. is below that which is generally regarded as necessary for success in a secondary course/

course. It is unlikely that any of these pupils will make a success of the secondary school at least as far as examinations are concerned. Only in the case of child K is there any improvement on the first year marks. There is a general fall in the marks after the first term of the first year. Indeed, the steady improvement in the case of the child K is remarkable and further investigation is necessary to find an explanation of this exception. Generally, the further these pupils advance in the course the lower do their marks become.

The difficulty in dealing with borderline cases may be emphasised by the following case. The pupil's rank in the December bursary examination was 55 - just below the pass level - and his I.Q. was 119. During the first term in a Central School his average mark was 72. In August, however, he enrolled in a secondary school and had an average mark of 55 for his first term examination. Apparently this mark deterred him from continuing the secondary school course and he was transferred back to his former class in the Central school. His subsequent marks were 71, 74, 73, 73, giving an average of 73 for the second year. This borderline pupil then, seemed to find the Central School course adapted to his abilities whereas his short spell in the Secondary School did not prove successful. On the other hand, once he had become adjusted to secondary school conditions he might have been quite successful there also. This seems to support/

support the argument for a probationary period of one year in the secondary school.

From these results it may be concluded that

(1) a single examination, particularly where the questions are of the essay type, does not give similar results to an intelligence test, scholastic test, Qualifying examination of a different type and teachers' estimates;

(2) the selection of pupils for secondary schools by means of a single examination is not reliable. Some pupils who gain awards do not justify their selection while others who fail to gain bursaries would likely profit by a secondary school course of instruction;

(3) difficulties of selection are most acute at the borderline.

XII. SUMMARY OF CONCLUSIONS

In this investigation the following are the main conclusions which have been drawn.

1. The most effective individual measure for predicting success in the first and second year terminal examinations in the secondary school is the English examination.
2. The predictive value of the teachers' estimates could be improved if some method of standardisation were adopted.
3. A weighted combination of the results of the Qualifying examination, the intelligence test and the teachers' estimates gives the best prediction of the first year terminal examination marks in the secondary school.
4. The marks in the first year examinations in the secondary school give a fairly reliable index of a pupil's marks in the second year examinations.
5. The Qualifying examination has a better predictive value for secondary school marks than the intelligence test.
6. The intelligence test picks out the less able children more efficiently than the Qualifying examination while the latter is more effective in selecting the clever children.
7. A single examination in English and Arithmetic, particularly where the questions are of the essay type, does not give similar results to an intelligence test, scholastic test, teachers' estimates or an examination of a different type.

8./

8. The selection of pupils for secondary schools by means of a single examination is not reliable. Some pupils who gain awards do not justify their selection while others who fail to gain bursaries would likely profit by a secondary school course of instruction.
9. Difficulties of selection are most acute at the borderline.

BIBLIOGRAPHY

BIBLIOGRAPHY

1. The Schools of Hellas. Kenneth Freeman.
   Macmillan & Co., Ltd. 1907.

2. The Meaning of National Examination System. Pan Q.
   China Critic. 1928, I, 512-515.

3. The Statistics of Examinations. Prof. Edgeworth.
   Journal of the Royal Statistical Society. Vol.51, 1888.

4. The Element of Chance in Competitive Examinations.
   Prof. Edgeworth.
   Journal of the Royal Statistical Society. Vol.53, 1890.

5. On Problems in Probabilities. Prof. Edgeworth.
   Philosophical Magazine. August, 1890.

6. Educational Reform. C. W. Eliot. New York, 1901.

7. Proceedings of the First Annual Convention of the College
   Association of the Middle States and Maryland. 1890.

8. The Examination Tangle and the Way Out.
   New Education Fellowship, 1935.

9. Conference on Examinations. Bureau of Publications,
   Teachers College, Columbia University, 1931.

10. An English Bibliography of Examinations 1900-1932.
    M. C. Champreys.
    International Institute Examinations Enquiry.
    Macmillan, 1934.

11. An Examination of Examinations. P. Hartog and E.C. Rhodes.
    International Institute Examinations Enquiry.
    Macmillan, 1935.

12. Free Place Examinations. Board of Education.
    Education Pamphlet No. 63, 1928.

13. The Marks of Examiners. P. Hartog and E.C. Rhodes.
    International Institute Examinations Enquiry.
    Macmillan, 1936.

14. La Correction des Epreuves. Escrites dans les Examens.
    Commission Française pour L'Enquête Carnegie.
    Paris, 1936.

15./

15. Schulerauslese. Kritik und Erfolge. O. Bobertag. Berlin, 1934.

16. Geist und Torheit auf Primanerbänken. W. Hartnacke and E. Wohlfahrt. Dresden, 1934.

17. Examinations and their Substitutes in the United States. I. L. Kandel. Bulletin 28. The Carnegie Foundation for the Advancement of Teaching, 1936.

18. The Intelligence of Scottish Children. Scottish Council for Research in Education. University of London Press, 1933.

19. The Prognostic Value of University Entrance Examinations in Scotland. Scottish Council for Research in Education. University of London Press, 1936.

20. Rise and Progress of Scottish Education. A. Morgan. Oliver and Boyd, Edinburgh, 1927.

21. A History of Secondary Education in Scotland. Clarendon Press, Oxford, 1909.

22. Scottish Education Department Circular 44. December, 1921.

23. The Year Book of Education, 1934. Lord Eustace Percy, M.P. Evans Bros., Ltd., London.

24. Code of Regulations for Day Schools in Scotland, Art. 15, 1923.

25. Report of the Advisory Council to the Scottish Education Department on the General Organisation of Day Schools and of Continuation Schools and Classes, 1923.

26. Committee on Education and Industry, Second Report, 1928.

27. The Education of the Adolescent. Board of Education Report of Consultative Committee. H. M. Stationery Office, London, 1926.

28. Essays on Examinations. International Institute Examinations Enquiry. Macmillan, 1936.

29. Prognostic Value of University Entrance Examinations in Scotland. University of London Press, 1936.

30./

30. Psychological Tests of Educable Capacity. Board of Education Report.
H. M. Stationery Office, 1924.

31. The Value of Intelligence Tests in Scholarship Examinations. R. R. Dobson.
The Forum of Education, Vol. I, No. 1, 1923.

32. A Statistical and Psychological Investigation of Intelligence Tests. D. W. Oates.
The Forum of Education, Vol. VI, No. 1, 1923.

33. The Nature and Validity of Subjective Estimates of Intelligence. D. W. Oates.
The Forum of Education, II, 1924.

34. The Reliability of Examinations. C. W. Valentine.
University of London Press, 1932.

35. Examination and Intelligence Test Forecasts of School Achievement. A. Donald Amos.
British Journal of Educational Psychology, 1931, Vol. I, Part 1.

36. Statistical Method in Educational Measurement. A. S. Otis.
Chicago, 1926.

37. The Predictive Value of Intelligence Tests for Secondary Education. J. W. Collier.
British Journal of Educational Psychology, February, 1933.

38. Discrepancies Between the Results of Intelligence Tests and Entrance Examinations to Secondary Schools. A.G.Hughes.
British Journal of Educational Psychology, Vol. IV, Part 3, 1934.

39. The Value of Intelligence Tests in an Examination for Selecting Pupils for Secondary Education. G.H. Thomson.
British Journal of Educational Psychology, Vol. VI, Part II, 1936.

40. The Necessary Imperfections of an Examination. F. Sandon.
British Journal of Educational Psychology, Vol. V, Part II, June 1935.

41. The Predictive Value of School Examinations and Psychological Tests. F. Sandon.
Transactions, British Association, 1933.
Sections J and I.

42./

42. Tests of Ability for Secondary School Courses. F.M. Earle. University of London Press, 1936.

43. Selection of Children for Secondary Education. Davies and Jones. Harrap, 1936.

44. Edinburgh Corporation, Education Committee. Report on Experiment in the Use of Standardised Tests, Conducted by Education Department, University of Edinburgh, October, 1934.

45. Measurement in Secondary Education. P. M. Symonds. Chap. XIX, Prediction of Success in High School. Macmillan, New York, 1927.

46. Comparison of Elementary and High School Grades. W.R. Miles. Iowa City, University of Iowa, 1911.

47. Educational Guidance. T. L. Kelley. New York, Teachers College, Columbia University, 1914.

48. Tests in Educational Guidance. W. M. Proctor. School and Society, Vol. 8, 1918.

49. A Study in Educational Prognosis. E. K. Fretwell. Teachers College Contributions to Education No. 90. Teachers College, Columbia University, New York, 1919.

50. The Validation of Intelligence Tests. A. M. Jordan. Journal of Educational Psychology, Vol. 14, 1923.

51. Correlations of Four Intelligence Tests with Grades. A. M. Jordan. Journal of Educational Psychology. Vol. 13, 1922.

52. Psychological Tests and Guidance of High School Pupils. W. M. Proctor. Journal of Educational Research. Monographs, No. 1. Public School Publishing Company, Bloomington, Illinois, 1923.

53. The Relation Between Grade School Record and High School Achievement. C. C. Ross. Teachers College Contributions to Education No. 166, New York, 1925.

54. Sectioning Junior High School Pupils by Tests and School Marks. F. D. Brooks. Journal of Educational Research, Vol. 12, 1925.

55./

55. A Detailed Analysis of Achievement in the High School.
C. W. Flemming.
Teachers College Contributions to Education No. 196,
Teachers College, Columbia University, New York, 1926.

56. The Psychology of Adolescence. F. D. Brooks.
Harrap and Co., Ltd.

57. Detailed Factors in Latin Prognosis. O. M. Clem.
Teachers College Contributions to Education No. 144,
Teachers College, Columbia University, New York, 1924.

58. A Study in Latin Prognosis. W. S. Allen.
Teachers College Contributions to Education No. 135,
Teachers College, Columbia University, New York, 1923.

59. Experimental Tests of Mathematical Ability and their Prognostic Value. A. L. Rogers.
Teachers College Contributions to Education No. 89,
Teachers College, Columbia University, New York, 1918.

60. Prognosis Tests in Modern Foreign Languages.
Publication of the American and Canadian Committees on Modern Languages.
Macmillan, New York, 1929.

61. The Prediction of Scholastic Success. R.D. Collmann and C. Jorgensen.
Education Research Series, No. 35.
Melbourne University Press, 1935.

62. The Entrance Examination in View of Later School Performances. Frits Wigforss.
Publications of the Swedish Society for a Psychological Pedagogical Institute, Stockholm, 1937.

63. The Standardisation of Group Tests and the Scatter of Intelligence Quotients. A Contribution to the Theory of Examining. G. H. Thomson.
British Journal of Educational Psychology, Vol. II, Parts I and II, 1932.

64. The Essential of Mental Measurement. Brown and Thomson.
Cambridge, 1925.

65. Statistics in Psychology and Education. H. E. Garrett.
Longman's Green & Co., 1937.

66./

66. Mental Tests. F. M. Freeman.
Harrap & Co., Ltd.

67. Measuring Intelligence. Terman and Merill.
Harrap & Co., Ltd., 1937.

68. On the Computation of Regression Equations, Partial Correlations, etc.. G. H. Thomson.
British Journal of Educational Psychology, Vol. 23, Part 1, 1932.

APPENDIX I.

FOLLOW-UP INVESTIGATIONS - TABLES OF RESULTS

APPENDIX I.

TABLE I.

Percentages Gaining School Certificate (Valentine).

Centre	S.C. in 4 years			S.C. in 5 years		
	Top	Middle	Bottom	Top	Middle	Bottom
A,B,C,E & F	50.5	52.2	42.4	74.8	80.5	74.8
H	92.3	92.3	84.6			
I	87.5	80.0	56.3			
J	50.0	35.6	22.2	75.7	62.6	52.8

TABLE II.

Percentages Gaining Matriculation (Valentine).

Centre	% gaining Matriculation in 4 or 5 years by		
	Top Third	Middle Third	Bottom Third
A,B,C,D & E	38.4	31.2	19.2
H	53.8	53.8	15.4
I	50.0	40.0	18.8
J	42.6	28.3	14.8

TABLE III.

Correlation Between Entrance Order and School Performance Four or Five Years Later (Valentine)

Centre	Correlation of Entrance Order with		
	S.C. order	order in school's own exams.	blended S.C. and school's own order
A,B,C,D,E	.03 (235)	.08 (221)	.06 (207)
H	.39 (24 boys)		
	.60 (15 girls)		
I	.44 (40)	.41 (40)	.42 (40)
J	.40 (625)	.39 (316)	.39 (316)

TABLE IV.

Distribution of Extraneous Handicaps (Hughes).

Handicapped by	I Winners			E Winners		
	Boys	Girls	Total	Boys	Girls	Total
Home	9	14	23	2	1	3
Health	10	8	18	5	4	9
Character	17	13	30	6	7	13
Attendance	2	2	4	0	1	1

TABLE V.

Home Conditions (Hughes)

Homes	I Winners			E Winners		
	Boys	Girls	Total	Boys	Girls	Total
Very favourable	4	2	6	2	6	8
Favourable	17	14	31	26	23	49
Tolerable	36	6	12	2	1	3
Unfavourable	3	8	11	0	0	0

TABLE VI.

Percentage No. of Children Reported as Suffering From Character Defects (Hughes)

	I Winners			E Winners		
	Boys	Girls	Total	Boys	Girls	Total
By Elementary Schools	57	43	50	20	23	22
By Secondary and Central Schools	55	30	42	17	12	14

TABLE VII.

Expectation of Teachers (Hughes)

Did you think the pupil would win a scholarship?	I Winners			E Winners		
	Boys	Girls	Total	Boys	Girls	Total
Yes, certainly	6	5	11	9	7	16
Yes, probably	6	6	12	14	11	25
Yes, possibly	10	10	20	5	9	14
Doubtful	6	0	6	2	0	2
No	2	8	10	0	3	3

TABLE VIII.

Correlation of First Year Average Grade in High School with Various Factors (from Kelley)

N = 59

7th grade average grade	.719
6th grade average grade	.728
5th grade average grade	.531
4th grade average grade	.624
Composite of elementary school grades	.789
Teachers' estimate of intellectual ability	.72
Teachers' estimate of conscientiousness	.62
Teachers' estimate of emotional interest in school work	.58
Teachers' estimate of oral expression	.63
Composite of teachers' estimate	.76
Composite of Mathematics, English and History tests	.51
Composite of all data	.89
Age	- .31

TABLE IX.

Correlation of Marks in First Year of Junior High School with Various Tests (from Fretwell)

	1916	1917
Thorndike Visual Vocabulary Scale A	.34	.33
Thorndike Reading Scale Alpha 2	.43	.40
Composition	.32	.35
Spelling	.32	.44
Trabue Completion B	.16	.22
Trabue Completion C	.31	.20
Woody Multiplication	.28	.33
Woody Division	.27	.30
Woodworth Wells Opposites	.26	.36
Woodworth Wells Easy Directions	.29	.23
Woodworth Wells Mixed Relations	.17	.17
Average of eleven Correlations	.29	.30
Composite of eleven tests	.57	.56
Marks in first six grades	.49	.49
Age	- .34	- .34

TABLE X.

Average Correlation Between Marks in the Elementary School and Marks in the First Year at High School (Ross)

	1916	1917	1918	1919
Reading		.35		
Spelling		.36		
Arithmetic		.52		
Geography		.50		
English		.59		
Fine Arts		.35		
History		.40		
Special Subjects		.26		
Deportment		.35		
Effort		.37		
Days Present		.00		
Simple Average of All Subjects		.60		
Weighted Average of Selected Subjects	.68	.67	.56	.65
Age		- .36		

TABLE XI.

Correlation of Marks in the Junior and Senior High School with Various Factors (Flemming)

	Junior High School	Senior High School
Terman Mental Group Test B	.60	.63
Miller Group Test A	.47	.49
Otis Higher Self-Administering Test B	.52	.32
Haggerty Reading Signia 3	.53	.30
Chronological Age	- .29	- .30
Estimates of		
Health	- .21	.35
Energy	.45	.34
Intelligence	.80	.70
Industry	.69	.50
School Attitude	.74	.50
Emotional Balance	.44	.31
Leadership	.51	.30
Will and Persistence	.72	.41
Conscientiousness	.61	.31
Desire to excell	.57	.46
Mean Teacher Rating	.70	.56

APPENDIX II.

RECORD CARD FOR SECONDARY SCHOOL CHILDREN

APPENDIX II.

Record Card for Secondary School Children

Name, ...

Year and Class, ... *Session* .....

		1st Term	2nd Term	3rd Term	Total.
Attendances.	Possible				
	Actual				
	Times Absent				
		Marks obtained in			
Subjects.	Full Value.	1st Term	2nd Term	3rd Term	Total.
English Lang. and Lit.	100				
History and Civics	50				
Geography	50				
Physics	50				
Chemistry	50				
Botany	50				
Zoology	50				
Mathematics	100				
French	100				
German	100				
Spanish	100				
Latin	100				
Greek	100				
Art	100				
Music (Leaving Certificate Subject) ...	100				
Domestic Subjects	100				
Economics	100				
Commercial Subjects	100				
Physical Training and Hygiene ...	Ex.				
Music (General)	Ex.				
Total Marks					
Possible Value					
Pupil's Average Percentage					
Number in Class					
Pupil's Place in Class					
Progress					
Conduct					

Signature of Parent or Guardian—1st Term

2nd Term

3rd Term