THE UNIVERSITY OF MANITOBA

A COMPARISON OF CERTAIN OBJECTIVE AND ESSAY-TYPE TESTS IN HISTORY

BEING A THESIS SUBMITTED TO THE COMMITTEE ON POSTGRADUATE STUDIES IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS

BY

MARY BELLE ELDER CANNON

BRANDON, MANITOBA

APRIL, 1935

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

LIST OF TABLES

LIST OF FIGURES

CHAPTER I.

INTRODUCTION

Purpose of the Study

It is the purpose of this thesis to make a comparative study of certain objective and essay-type tests in history and to discover, if possible, which is the more valuable for use at the grade eight level in the schools of Manitoba.

In justification of the study, expert opinion as to the value of a test as an instrument for measuring results and for directing methods of teaching has been sought among the most modern writers in this field. The books, publications and authors consulted are listed in the accompanying bibliography.

In measuring the results obtained from actual tests submitted, statistical methods have been employed and each step carefully reported. The tests, as submitted to the pupils, are herewith attached in Appendix A; the scores obtained in each test are reproduced in Appendix B; the correlation tables and tables of findings produced by the study are presented in Appendix C.

In conclusion, an attempt is made to interpret the findings obtained and to indicate how, even under existing difficult conditions, testing in History at the grade eight level may be done throughout Manitoba, in a way which will do greater justice to the pupils and at the same time operate to improve methods of teaching.

CHAPTER II.

Old and New-Type Tests: Arguments for and Against Each, as Expressed by Authoritative Writers.

The word "test" implies the purpose for which examinations are given in school. The achievement of pupils in school is tested in order that the strength of the foundation may be ascertained before we proceed to build upon it. The discussion in this thesis will centre about the question of whether the testing of school achievement with particular reference to History has been done or is being done in a way that truly reveals understanding of the subject.

In 1927, a committee consisting of Ernest Horn, John Bassett, Henry Johnson and others, under the chairmanship of A. C. Krey, submitted the following report to the Council of the American Historical Association; "The old-fashioned examination has been under fire for some time, both as a device of instruction and as an accurate measure of achievement. More important is the fact that psychologists have devised methods of testing and measuring results which promise to prove useful in the teaching of the Social Studies. The need for the most accurate testing device possible arises in connection with nearly every division of this study. While relying still upon sensitive and well-trained

teachers to furnish the most important gauge of results, it is desirable to enlist whatever aid the new tests may offer. There is need of such help in the determination of objectives, organization of content, and grade-placement as well as in the formulation of examinations. For these reasons, it seems desirable to make the improvement of tests in the social studies a major issue."[1]

In their book entitled "Objective Tests,"[2] Orleans and Sealy write: "A test is not adequate if it merely gives a percent mark indicating that the pupil passes or fails, and that reaches its final resting place and usefulness in the teacher's record-book. The test results should be of such a nature that the teacher can determine for her own information, or for that of the teacher who is to have charge of the pupil the following year, what his level of achievement is in comparison with the standard levels of achievement for the school system. The test results should also enable the teacher to compare the product of her efforts with that of other teachers in the school system and the relative efficiency of her teaching of the several school subjects."

---

1 Historical Outlook, Vol. XVIII., 1927; pp. 119-120.

2 Orleans and Sealy, Objective Tests; World Book Co., Yonkers, 1928, p. 2.

The Function of Examinations

Ruch reports[1] that four functions have been claimed for examinations, namely, (1) motivation of the learning of the pupils, (2) maintenance of standards of accomplishment, (3) training in the use of the English language and (4) measurement of accomplishment. The fourth of these is the function commonly in the mind of teacher and parent, nevertheless the others are worthy of brief consideration.

To quote Ruch, "If examinations are of value as motivation, there are certain qualities which it is well that they should have. In the first place, they must be held in esteem by the pupils; the examination which can be met by a few weeks' cram, to be forgotten within the next few weeks, cannot fail to produce an undesirable attitude among students. Similarly, the examination which is sufficiently subjective to permit teachers' feelings toward the student to color the marking will inevitably result in the student doubting its justice. In the second place, if examinations are to be of value for motivation, they must come at frequent intervals, in order that the pupil may be able to keep and to read a record of his progress. Thirdly, they must be looked upon as guidance,

---

[1] Ruch, G. M. The Objective or New-Type Examination; Scott, Foresman & Co., Chicago, Ill., 1929, p. 9.

rather than as a trial at which he must prove his case or be condemned."

If examinations have for one function the maintenance of standards of achievement it would seem to be imperative that they be applied rather frequently, and possess a high degree of accuracy as a measuring device. If they can serve this purpose it is obviously of special value in communities like our own where so many pupils are receiving their schooling in one-roomed rural schools with practically no assistance from supervisors. For such a purpose, tests would necessarily confine themselves to essential points and would parallel the curricular units with exactness; they would, of necessity, be the type to discourage cramming or specific coaching.

The third function claimed, namely, a training in English, Ruch denies.[1] "Most teachers," he says, "have noted that final examinations show a quality of diction, grammar and spelling markedly inferior to the products of the regular English classes. Some teachers are convinced that the written examination has a negative value in English. Actual conditions of the examination period are unfavorable to good linguistic expression. Again, the pupil realizes that his paper will be graded on facts, not style, and will govern himself accordingly; and lastly, it is unlikely

---

[1] Ibid. p. 14.

that language habits, in themselves complex, will arise as by-products of frenzied efforts at setting down facts in limited time. In any case, even allowing that examinations might possess some value as training in English, the fact remains that examinations are a poor substitute for a lesson period in English; examinations are obviously not intended as 'training' in anything."

We arrive, then, at the fourth function; i.e., the measurement of achievement, which has always been, admittedly, the principal reason for examinations. If the examination does this, the by-products such as motivation, language-training, training in critical judgment, etc., may be ignored so far as the test-maker is concerned.

Forms of Testing

The oral test, which was the first form of testing, and which, indeed, kept its popularity until the time of Horace Mann, will not be discussed at any length. The teacher's daily questioning is the best form of oral teaching, and may reveal to her more about her pupils' achievements and attitudes than the final written examination. In this connection it may be stated here, that the rating of the pupils according to teachers' judgments, which forms one basis of comparison in this study, was presumably founded largely on the results of oral testing in the class-room day by day.

The logical purpose of the oral test is, moreover, instructional; a record of the measurement obtained through it could not be thoroughly reliable if that were the sole means of testing. So many factors, such as nervousness, limited time, limited sampling, the fact that a question, answered orally by one pupil, is spoiled for the rest of the class, - enter into the oral test, that its value for measurement is subordinate to its value as part of initial instruction.

Accordingly, the discussion in this thesis will centre about the familiar and widely-used forms of testing, namely the traditional, or essay-type, test and the new-type or objective test.

Orleans and Sealy[1] say that in order that an examination may be a real test of achievement in school work, it should satisfy the following requirements.

(1) It should be representative of the subject-matter, the mastery of which it is intended to measure.

(2) It should provide a comprehensive measure of the achievement of each pupil.

(3) The scoring should be objective so that different teachers will assign the same mark to any answer.

(4) The wording of the questions should lead pupils to answer in forms that can be marked objectively.

---

[1] Orleans and Sealy, op. cit. p. 6.

(5) The amount of written expression should not be so much that it will interfere with the pupil's ability to show that he can answer the questions.

(6) The standards by means of which pupils' test-scores are interpreted should not be determined arbitrarily, but should be based upon the average achievement of the pupils in the several school grades.

(7) The test scores should be in such form that they can be used for a large number of purposes and applied efficiently to improve the school product.

Professor Howard C. Hill says[1] that a test should (1) acquaint the teacher with the needs and capacities of his pupils; (2) aid the teacher in selecting materials for study and in planning study activities; (3) inform the teacher whether the learning outcomes sought in a course have been realized, and (4) help the teacher to diagnose pupil difficulties and to formulate remedial measures for the overcoming of intellectual obstacles.

Ruch,[2] giving the criteria of a good test, or examination, enumerates them as (1) validity; that is, the test must measure what it purports to measure; (2) reliability; that is, the degree of accuracy with which the examination measures whatever it does measure; (3) ease of administration and scoring; (4) standards for evaluation of

---

[1] Hill, Howard C. The Use of Tests in the Teaching of the Social Studies; Historical Outlook, Vol. XX., January, 1929, p. 87.

[2] Ruch, G.M., op. cit. p. 27.

results; (5) availability of duplicate forms.

This discussion will follow the lines of Ruch's criteria, with particular reference to history-testing.

Validity of a Test

As early as 1845 Horace Mann championed the development of some more objective form of examinations to replace the traditional type. The movement has been given its greatest impetus following on the investigations of Starch and others. The principal objections to the essay-type test which have arisen as a result of these investigations reduce to the questions of validity and subjectivity of scoring. Ruch lists them as (1) subjectivity of scoring which lowers reliability, (2) the sampling must be limited to a small number of broad questions, thus reducing validity, (3) the time required to write lengthy answers is excessive, and (4) these examinations encourage bluffing.

What steps must be taken then to give the test-maker reasonable assurance that his test is valid, and that it does reveal truly the degree to which the aim of curriculum-makers has been achieved?

The first criterion of validity is "the judgment of competent persons."[1] The pooled judgment of a fairly large number of persons competent to judge on any subject

---

[1] Ibid. p. 29.

has been accepted by our law-courts for many years.

The second method of validating a test is by analysis of text-books. Assuming that the text-book has been selected for its value as, after the teacher, the most important agent in the accomplishment of the stated aim of the curriculum-makers, it is logical to assume that a careful analysis of it will disclose points of emphasis and guide the test maker in evaluating items.

Among the scientific methods of validating a test are: studies of the most frequently recurring errors; computation of the percentages of pupils answering each item correctly, or correlation with some established criterion. Ruch says that for informal class-room tests, "The judgment of one or two teachers will usually accomplish a degree of refinement commensurate with the needs of the average test. One thing, however, should be kept in mind; viz., long tests may be expected to be more valid than short tests and if a test is long enough, it will usually yield a reasonably valid measure, even though some of the items have very limited validity."[1]

Ballard defends"the new examiner" thus:[2] "The main advantage of a large number of questions is that they

---

[1] Ibid. p. 39.

[2] Ballard, P. B. The New Examiner; Hodder & Stoughton, London, England, 1929, p. 69.

cover a wide field. Instead of probing the mind at a dozen points, they probe it at a hundred points. The old examiner dips his hand a few times into the store-house of the pupil's mind and brings up samples which he assumes to be fairly representative of the whole stock. But chance has a hand in it as well, and the less frequently the examiner dips, the more active is the hand of chance. When in a particular instance, the examiner has taken a bad sample the candidate complains of his 'rotten luck.' The less 'luck' there is, the better; and the more questions are squeezed in the more luck is squeezed out. Thus the new examination gives the candidate a greater feeling of satisfaction than the old; a feeling that, at any rate, he got his deserts - that the examiner has put his finger on his strong points as well as on his weak points."

Oliver B. Floyd attacks the new-type, as well as the old-type test on the ground of validity.[1] He says, "Do the newer tests provide any better means of determining the absence or presence of the desired learning products where these objectives are attitudes, concepts, or ideals? Many more questions may be placed on examination papers, but the emphasis is still on subject matter."

---

[1] Floyd, Oliver, Validity of Tests in the Social Studies; Historical Outlook, Vol. XX., 1929, p. 11.

Dr. Hill, in an article in the same publication[1] gives a very adequate reply to Floyd's criticism. "Can a sense of evidence be taught?" he asks. "Can it be tested? If by a 'sense of evidence' is meant the ability to discover sources of reliable information; to weigh simple testimony, to evaluate varying degrees of credibility, and to differentiate between fact, opinion, supposition and hearsay, both questions can be answered in the affirmative." And he supports his arguments by concrete examples.

C. W. Odell replies to the contention that objective tests lack validity because they measure only memorized material and facts, by saying that it is not the desire of test-makers to abolish all essay-type questions from examination papers.[2] Ruch supports this statement when he says, "The traditional examination should be employed principally when the subject matter does not lend itself to completely objective measurement; even in such cases the results must be taken with a good deal of conservatism. A combination of traditional and new-type examinations should probably be used in any school subjects, especially where present knowledge is unable to suggest purely

---

[1] Historical Outlook, Vol. XX., p. 7 - 9.

[2] Odell, C. W. Educational Measurement in High Schools; The Century Book Co., New York, N. Y., 1930, pp. 12-16.

objective types of measurement."[1]

Odell, moreover, defends the testing of actual facts in history.[2] "It has been said," he writes, "that the factual side of history is not important. This, however, is not true. Before an individual can think in historical terms he must possess the proper data upon which to base his thinking. Moreover, for purposes of general culture, understanding of allusions in literature, etc., it is important that one be in possession of a fair stock of historical facts."

Reliability of a Test

Second only to validity as a criterion of the worth of a test, is reliability, in itself really one aspect of validity. A valid test is necessarily reliable, although a reliable test is not necessarily valid. Reliability refers to the degree of accuracy to which a test measures whatever it does measure, not necessarily what it is claimed to measure.

Objectivity is a prime essential for reliability of measurement. Measurement implies accuracy. Examinations are measuring instruments, then, only in so far as personal opinions, temperament, or whims fail to affect the reading of the instrument.

---

[1] Ruch, G. M. Op. cit. p. 277.
[2] Odell, C. W. Op. cit. p. 277.

Numerous studies have been made on the general problem of the reliability of school marks and marking systems. A survey of a selected few of these will be enlightening.

F. J. Kelly[1] reports a study made by Professor Johnson of marks in the University of Chicago High School for the purpose of comparing the marks of the several departments of that school. The standards of the various departments are obviously widely varied. English teachers fail almost three times as many pupils as do Domestic Science teachers, and give about half as many A's. According to Johnson's table a student has approximately twice the chance of getting an A in German as he has in French.

Carl E. Hendrickson, reporting a similar study from Los Angeles[2] shows that the A's in Art are three times as many as in English, approximately twice as many as in Physical Education, and about two-thirds as many as in Music. He comments also, "It is probably significant to point out that 56% of the total marks were A or B or college recommending, and only 18% were D and E. In other words, the distribution is skewed over toward the highest marks considerably. Here it may be of interest to add

---

[1] Kelly, F. J. Teachers' Marks; Teachers College Contributions to Education, No. 66. Columbia University, New York, 1914, p. 11.

[2] Hendrickson, Carl E. School Marks at Van Nieys High School. Educational Research Bulletin, Los Angeles, Vol. VII., December, 1927, pp. 8-9.

that all the mental and standardized educational tests given in this school result in approximation to the normal curve. Also the level of intelligence here is about normal."

Ruch reports finding the same skewing upward in a study he made of 659 schools for a six-week period in the University of Oregon High School, Eugene, Oregon.[1]

These three studies show the same general tendencies. While letter-grading is admittedly a very loose grading and each school must settle its own policy regarding it, yet the normal curve should suggest the relative proportions of letter-grades in contrast with the absolute numbers, and it is highly improbable that in an unselected group, the percentage of A's should be 46.8 while the E's are 4.5. It is much more reasonable to suppose that the variation is to be explained by such facts as non-adherence to school standards; differences in the subjective standards of different teachers, etc., than by such facts as differences in efficiency of instruction or in special ability of pupils.

---

[1] Ruch, G. M. Improvement of the Written Examination; Scott, Foresman & Co., Chicago, Ill., 1924, pp. 47-48.

Dearborn,[1] Johnson,[2] Starch and Elliott[3] and others, have made studies which show that practically all ordinary school marks given to pupils are highly subjective; that is, that they depend to a considerable degree upon the person giving them and vary greatly if given by different individuals at different times. Starch and Elliott report having submitted exact copies of the same examination papers to a large number of teachers. One hundred and forty-two English teachers marked the same English paper and the marks varied from 50 to 98 with not more than two agreeing. One hundred and fifteen teachers graded the same paper in geometry, and contrary to the common belief that uniformity of marking is easy to secure in Mathematics, the marks varied from 28 to 92.

B. D. Wood reports a study of college entrance papers in Algebra and Geometry that were scored independently by two different teachers. The results show that if about 30% of the candidates are failed by each reader, the chances are that less than 60% of those failed by

---

1 Dearborn, W. F. School and University Grades; Bulletin of University of Wisconsin, No. 368 High School Series No. 9, Madison, Wis., June, 1910, p. 59.

2 Johnson, F. W. A Study of High School Marks; School Review, Vol. XIX., January, 1911, pp. 13-24.

3 Starch and Elliott. School Review; Vol. XX., pp.442-57; Vol. XXI., pp. 254-59; Vol. XXVI. pp. 676-81.

Reader No. I will be the same as those failed by Reader No. 2. In other words, more than 40% of those failed by one were passed by the other.[1]

Ruch, DeGraff, Gordon and Others, conducted a series of tests to determine the reliability of the mark given by teachers on essay-type examinations for eighth grade pupils.[2] Their findings are as follows.

(1) Sixteen eighth-grade diploma examinations representing eleven different states yielded average reliability coefficients as follows:

(a) .62 when the examination was constant and the reader was variable;

(b) .43 when the reader was constant and the examination was variable;

(c) .38 when both reader and examination were variable;

(2) These agreements are from 7½% to 20% higher than chance assignments of marks;

(3) The unreliability of examinations is due to two general causes; viz: limited sampling and subjectivity of scoring;

(4) Unreliability due to subjectivity is completely or almost completely eliminable;

(5) Unreliability due to limited sampling may be minimized by the use of examinations composed of 100 to 250 short questions rather than of 5 to 10 broad questions; i.e., by adopting an extensive plan of sampling rather than an intensive one;

---

1 Wood, B. D. Measurement in Higher Education, World Book Co., Yonkers, N. Y., 1923, pp. 124-25.

2 Ruch, DeGraff, Gordon and Others. Examination Methods in Social Studies; Scott, Foresman & Co., Chicago, Ill., 1926, pp. 6-21.

(6) Differences in the average marks assigned a set of papers by two independent scorers may be as great as 25%; the central tendency of such differences being found to be about 8 to 10 points. Such differences were greatest when both the scorer and the examination were variable from year to year;

(7) The differences argue against the accuracy of the 100 per cent grading scale;

(8) Large differences in the variability in the distribution assigned by different scorers to the same papers or by the same scorer to papers of different years were found making a mark of 80% on one examination statistically equivalent to a mark of 65% or of 95% on the examination on the same subject the next year.

How then is reliability to be secured for a test? Reliability, we say, is that phase of validity which refers to the accurace of a test as a measuring device; it is presupposed when validity is established; it assures the stability of scores for the same individual when equally difficult and similar tests are given. It is guaranteed in two principal ways, according to Ruch;[1] viz., (1) objectivity of scoring, and (2) extended sampling. When personal opinion can affect the evaluation of the paper, reliability is lowered , but personal opinion can be almost entirely eliminated by use of a test where the answer is either right or wrong, that is by the use of the new-type or objective test; and it can be greatly lessened by lengthening the test and sampling extensively.

---

[1] Ruch, G. M. The Objective or New-Type Examination; op. cit. p. 62.

Administration and Scoring of a Test

The third standard on which a test is evaluated is the degree of difficulty it presents in administration and scoring. Ease of administration is judged from two points of view; first and most important, the clarity of instructions to the pupil, and second, the clarity of instructions to the examiner. It is a question of the care with which the test is constructed. The essay-type test, of course, simply gives a number of essay topics, and the pupils interpret them and discuss them, each in his own way. The objective test requires vastly more time for its construction. The wording of each item is a matter of most critical consideration. Explicit instructions to the examiner accompany the test and directions to the pupil are printed, together with samples showing how he is to indicate his answers, on the test itself or on each unit of the test, if it is in sections. The examiner experiences no difficulty in administration, the test-maker considers it as part of his job to make the test "fool-proof." Ruch and Stothard, in discussing the building of a test, say,[1] "A test whose administration cannot be learned by the average teacher in an hour or two

---

[1] Ruch and Stothard. Tests in High School Instruction; World Book Co., Yonkers-on-Hudson, N. Y., 1927, p. 57.

is not likely to succeed. At the same time the beginning examiner must be warned about a multiplicity of small but important details, such as instructions covering the distribution of blanks, the filling-in of pupil information data, the observance of time-limits; the prevention of disturbing factors and other requirements of good test conditions."

When we come to consider the question of scoring we touch the point upon which the new-type test has most often been granted superiority.

There are two aspects to the question of ease of scoring; first, economy of time and effort, and, second the more mechanical the scoring, the more objective the results. Ruch declares[1] that there is no dispute on the point of economy of time and effort. Objective tests, if well planned as to mechanical features, may be scored from two to five times as rapidly as can essay tests of comparable length. A ten-question essay-type test he compares with a forty to seventy-five item objective test and he says that the latter can be scored twenty or thirty more per hour.

---

[1] Ruch, G. M. Objective or New-Type Tests, op.cit. pp.118-19.

Ballard supports the objective tests on another count in regard to scoring.[1] He advocates having the pupils mark the papers; a perfectly sound idea since the scoring is done from a key. He contends that by so doing, the teacher has converted the examination into a lesson and that this is one of its main advantages.

At this point it may be mentioned that in general, teachers during their preparatory courses in high school, university and normal school, spend vastly more time in fitting themselves to present well and truthfully the subjects they propose to teach than they do in fitting themselves to test the results of their teaching. It would seem that the general opinion, both of the embryo teacher and of those responsible for the professional training of the same, is, that it is much more important that he give evidence of his own knowledge than that he prove his adequacy to test the results of his efforts to impart it to others. In other lines of practical and experimental work the accuracy of testing and the intelligent interpretation of the results obtained is the final proof of the work. It would seem reasonable that a

---

[1] Ballard, P. B. The New Examiner; op. cit. p. 135.

thorough training in the methods of testing ought to be one of the most important phases of the work of the teacher-training and normal schools everywhere.

Standards and Duplicate Tests

The fourth criterion of a good test is that it shall have a norm or standard for evaluation of results. The central feature of testing, from the standpoint of the teacher, is the reference to a standard of pupil performance. The 50% passing mark which is commonly accepted means little or nothing. Orleans and Sealy say,[1] "Since those who construct a test are not capable of determining its difficulty before it is given, a predetermined passing mark is eminently unfair. This unfairness is in practice, counteracted in part by lenient marking if the test is hard, or by severe marking if the test is easy. But such subjective scoring is well-nigh worthless and meaningless. A fair standard for a test can be assigned only after the test has been given and its difficulty determined from the pupils' answers. This is utterly opposed to traditional practice."

---

[1] Orleans and Sealy, op. cit. p. 30.

The last point in the discussion is the value and the possibility of constructing duplicate forms of tests. The usual practice of building tests anew each year, if they contain only five or ten questions, can have only one of two results; either there will be large differences in the difficulties of the examinations from year to year, or they will lead to the familiar practice of preparing elaborate notes (probably dictated) on the "important topics." Differences in difficulty of examinations make it almost impossible to tell whether successive classes differ in average ability or whether these differences are apparent rather than real. When objective tests are prepared, two or more equivalent tests can be constructed by increasing the number of items.[1] Well constructed duplicate examinations are sufficiently equal in difficulty to enable the teacher, by keeping records, to compare the performance of her classes year by year, and thus finally to arrive at a working norm. By using duplicate forms, too, she can be assured of grading pupils over a period of years by the same scale of marking.

---

[1] Ruch, G. M. Old and New-Type Examinations, op.cit.p.164.

Strength and Weakness of Objective and Essay-Type Tests

There are, of course, objections raised to the new-type examinations. It is claimed that they measure only memorized material and isolated facts; that they are difficult to construct; that they encourage guessing and bluffing; that they confuse pupils; that they mechanize education; and that after all the judgment of competent persons is more reliable. To all of which the reply must be, that, in the first place, their most ardent advocate does not wish to eliminate entirely the essay-type question; it is desired only to keep it in the place where it has a real value, as supplementary to the objective test; that the time spent on the construction of a test of the new type is more than repaid by the time saved in scoring, and in the accuracy of results. If the examinee guesses, or bluffs, it is not that he did not always do so; the difference is that with the objective test he is caught and is made to pay for it. The confusion to students is in reality only a consequence of pinning them down to

correct or incorrect answers, and there is no justification for failing to teach them to make that distinction. Odell claims that rather than mechanize education, the new examination "has revealed and emphasized differences in the capacities, characteristics, and needs of individuals which were scarcely at all realized before."[1] In response to the claim that agreement in judgment of competent persons is a more reliable guide in assigning a mark to a pupil, Boyd H. Bode declares that such agreement is impossible and need not be expected,[2] a fact which is fully demonstrated in the studies already quoted.

Charles E. Chadsey in the preface to Odell's book on Educational Measurement writes,[3] "The value of educational measurement as a means of determining the actual accomplishment of students in given types of academic work, through standardized tests of one kind or another, has been recognized for a number of years. While one does not find so many controversial articles with reference to the value or lack of value of educational measurement as he did a few years ago, the number of school systems

---

[1] Odell, C. W. op. cit. pp. 277-290.

[2] Bode, Boyd H. Journal of Educational Research, Vol. X., p. 179.

[3] Odell, C. W. op. cit. preface.

making a systematic use of various types of tests is steadily increasing.

"It is quite evident that the school system which makes a free and intelligent use of measurement from both the diagnostic and the prognostic points of view is bound to be carrying on a much more successful type of teaching than the school system which ignores almost completely the existence of these tests and their very genuine significance in education."

Lewis Terman in the preface to Ruch and Stothard's "Tests in High School Instruction"[1] writes, "In the Junior and Senior High School, pupil guidance, both educational and vocational, assumes outstanding importance. The student stands at the threshold of life, subject to conflicting desires and uncertain as to his abilities and as to the direction he should take. He is facing decisions which will affect his entire life; whether to leave school and go to work; whether to complete the high-school course; whether to prepare for college; whether to fit himself to enter one or other of

---

[1] Ruch and Stothard. Op. cit. Editor's Note.

the professions. Without objective information of the kind that is obtainable only by standard tests, the guidance of such a student can rest upon but little more than guess work. One is tempted to put it more strongly and to say that educational guidance without educational testing is professional quackery as much as in the case of a physician who refuses to employ the approved laboratory technique in the diagnosis and treatment of certain diseases."

A. C. Krey, answering the critics of the new-type test in the Social Studies says,[1] "The question of the relative reliability of instructors' judgments and the results of objective tests is one usually raised in discussion concerning the possibility of testing students' knowledge of relationship of facts rather than his memory of the facts themselves. Our experience with this question has been sufficient to justify some tentative observations. On a single test a comparison of the objective type and the essay-type seems to show that the objective-type offers a more accurate measure of whatever it measures, than does the teacher's judgment of an essay-type examination. The difference is less marked

---

[1] Historical Outlook, Vol. XIX., p. 160.

at the extremes, but with the great mass neither strikingly good nor strikingly bad, the objective type offers a means of definite distribution, whereas the essay-type too often leaves a large undistributed middle. Sometimes, too, where experiment has been made, the same papers in that middle group have been given different grades at different times by the same instructor."

From the pen of Krey comes one of the broadest and most unbiassed articles yet submitted on this very controversial subject.[1] In summarizing the conclusions which his experience seems to warrant, Krey says:

(1) "The objective-type of examination permits the testing of a much wider range of information than the essay-type. This is distinctly valuable in history and offers a vast improvement over the 'omnibus' questions."

(2) "The objective-type examination is more easily graded than the essay-type. In dealing with large classes under the somewhat hectic conditions which prevail at the close of the term, this is also a very important advantage."

(3) "The objective-type of examination yields a wider distribution of grades and therefore makes possible the determination of relative ranking in large classes more accurately than the essay-type."

"But," he adds, "against these advantages which are very real, the following disadvantages have appeared:

(1) The objective-type of examination requires much greater time in preparation. It is almost

---

[1] Krey, A. C. Historical Outlook, Vol. XVIII., p.119.

essential to have a group of teachers collaborate in preparing these examinations."

(2) "In order to avoid ambiguities and have questions require single precise answers, there is a marked tendency to limit the examination to matter which can be so treated. In other words, it seems an axiom that the greater the care in safe-guarding the questions, the greater the emphasis upon mere factual information."

(3) "The foregoing indictment seems to have been recognised by the students and it is already a campus tradition that the best preparation for an impending test of the objective type is to commit to memory tables of information and indices of text-books."

(4) "The actual cost of making and using the objective type of examination is much greater than the essay-type."

(5) "The test-maker, despite all efforts, has been unable to frame examinations which do not require relatively expert readers. The promise that such papers might be turned over to clerical assistants has proved vain, thus far.

(6) "Further, we have an uneasy feeling that in using these objective tests, most of the thinking has been done by the instructors, leaving to the student merely the task of applying a few bits of information involved in that thinking. We do feel that the student should have latitude to show how well he understands the condition of social phenomena. The objective test does not offer that latitude."

(7) "Another value which we have thus far been unable to test in this fashion is the students' ability to check impulsive judgments by thoughtful consideration of the various points of view which enter into any problem. There is a natural tendency to arrive at important social judgments on the basis of what might be termed purely emotional reactions, attitudes

created by early environment, accidental and inadequate contacts or temperamental "set." History, and all of the social sciences, clearly have as one of their chief functions, the task of safe-guarding society against the dangers of such judgments, by trying to establish the habit of circumspection both of fact and point of view. Thus far we have found no better way of attaining this end than by presenting a problem to the student and giving him free range to discuss it in his own way."

(8) "Still another element which seems to us important, and yet, up to this time, beyond the reach of the objective test is the very quality which by title the objective test seems to scorn, that is, the subjective values. Encyclopedic knowledge is of advantage of course, but we feel that effectiveness in presentation, orally or in writing, is likewise a very important element. We are inclined therefore to value somewhat more highly the student who knows how to make effective use of his information even if his range of information is not quite so great as that of some other less able to make use of it. Plagued by the thought that we may at some time be called upon to recommend the student for a position, academic or otherwise, we feel that we ought to know about these 'subjective qualities.' Thus far, we have not been able to gain any help from the objective tests in arriving at such opinions.

"For these reasons, the members of our department who use the objective-type test do not emplpy it for more than one-half the examination."

Professor R. M. Tryon, writing in 1927, stated:[1]

"In spite of all its defects, thousands of teachers of the social studies still make the traditional written

---

[1] Tryon, R. M. Historical Outlook, Vol. XVII., 1927, p.172.

examination their chief tool for determining the results of their teaching. Until scientifically determined means have been provided, teachers will continue to use traditional ones for testing their students. The present status of the written examination in the social studies would seem to indicate that improvement rather than abolishment is what is needed. Even the most ardent supporters of the standardized and new-type tests still feel that there is value in the traditional written examination if it is discriminately and judiciously used.

"It seems," Tryon continues, "that certain obstacles have arisen which the testing movement in the social studies has been unable to overcome. Specifically stated, some of these are:

(1) "When a test is once seen and used by a class its value for subsequent use is greatly diminished if not entirely destroyed.

(2) "To date, no one has succeeded in making a high-grade test in history which is as readily given and as easily graded as is demanded of tests in general.

(3) "The demand that history tests, in so far as possible be independent of attainments in other subjects, has been difficult to meet.

(4) "History-test makers have had difficulty in devising exercises that test a variety of mental processes. Too many test memory only.

(5) "To meet the demand that each exercise in a test must require an unequivocal answer has not always been easy, when faculties like reasoning and comparison are included."

Conclusion

Having in mind then, this agreement of competent opinion that history testing presents a difficult problem, and also the fact that a great many of the Grade VIII. teachers of Manitoba are teaching and testing in rural schools without the aid and guidance of experienced supervisors, the writer conducted an experiment in the testing of Canadian History at the Grade VIII. level for Manitoba pupils. The method of procedure, the results and a critical comparison of those results follow.

CHAPTER III.

Building, Administering and Scoring History Tests, Suitable for a Student Seeking Entrance to Grade IX. in Manitoba

In building a test, it is safe to assume, the test-maker desires to make a measuring instrument whereby the achievement of pupils at a certain level may be measured, duly recorded and compared with others at the same level. To-day an examination board has replaced the one-man test-builder of former years, and if the tests are built in committee, their validity is increased in proportion to the number of expert judges who have approved them.[1]

If, however, no further steps are taken to insure validity, or to test the test to ascertain whether it be valid, it may still lack that very necessary quality. Teachers are notoriously conservative and a group of two or three may easily be found who would agree that the traditional test needed no experimental evidence to testify to its worth. In most instances that is what happens when the essay-type test is constructed.

---

[1] Ruch, G. M. The Objective or New-Type Examination: op. cit. pp. 150-160.

Building, Administering and Scoring an Essay-type Test

In building the essay-type test for this experiment the writer examined the examination papers sent out by the Manitoba Department of Education for testing Grade VIII. Canadian History, between the years 1920 and 1932. The average number of questions was five, usually with one or two options; therefore it was decided to make a paper of five questions with an optional question. Then from those examination papers were chosen questions that, variously worded, had appeared on at least one-half of those papers. Some of the questions had appeared on almost every one of the papers, with only a slight alteration in wording. The six questions having the greatest frequency were selected and in the wording that seemed to the writer to be most clear, were compiled into one test-paper. The thought was not to make the best possible essay-type test, but to have for experimental purposes the most typical test used for this purpose in the province of Manitoba.

Having compiled the test, the writer asked a group of three teachers who have been closely in touch with Grade VIII. history-teaching for not less than five years, to approve or disapprove of it. The test, approved by

them, appears as Test B in Appendix A, Page xvi. Asked upon what grounds they based their approval, this group of teachers said that (1) in their opinion the several topics were of equal value, thus justifying the option, and (2) the topics represented various phases of Canadian History, each of which was important. The possibility of the pupil's memorizing "notes" on topics so frequently given was granted, but was considered justifiable on the grounds of their importance. The student had, it was claimed, no means of knowing which topics would be chosen out of a possible ten or twelve. If he were familiar with all the topics he had mastered his history. The element of luck was conceded but it was felt that it was an element in all examinations.

The test as compiled was accepted, therefore, and used in the experiment herein reported.

The administration of the test was very simple. The questions were written on the blackboard and the classes were given an hour and a half in which to answer them.

The scoring was done by five different markers, each one evaluating the answers according to his best judgment.

In preparing the objective test covering the same work and given for the same purpose, the writer followed the procedure outlined by Ruch in his book "Objective or New-type Examinations."[1]

Building, Administering and Scoring an Objective Test

An analysis of the text-book was made to determine the space allotted to each of the five sections into which it is divided, and the approximate weight given to the various types of history contained in it, namely: social, political, military, economic, exploration and settlement.

From this analysis a table of specifications was drawn up similar to that in Ruch.[2]

Based on this table of specifications, two hundred questions were prepared, as it was desired to have, finally, a test of one hundred and fifty items. This

---

1 Ibid. pp. 150-160.

2 Ibid. p. 151.

would allow for discarding upon revision.

The questions were written on filing cards, and were of three types, matching, true-false, and multiple choice. It was felt that a matching-test was valuable for testing the student's ability to associate important events with dates; and events with names of men closely associated with them.

No definite effort was made to construct a fixed number of true-false or of multiple-choice items. As Ruch suggests,[1] some items expressed themselves readily in the one form, some in the other.

In making up the true-false test, the decision concerning whether the statement should be true or false was left to chance, each time on the toss of a coin, so that there is no system of arrangement.

In making up the multiple-choice test, the five-answer type of item was used so that correction for guessing might be avoided in the scoring. In order that a strip key might be used in scoring, the blanks were at the end of the item, and the responses were indicated by numbers.

To insure validity, the two hundred questions thus

---

[1] Ibid. p. 155.

prepared were then given to six experienced history teachers in turn. These teachers have been in close touch with grade eight history in Manitoba for some years, and three of them have had special training in history. The following request accompanied the test. "In the first vacant space on the back of the card, indicate your opinion of the question this: if perfectly satisfactory, mark I.; if fairly satisfactory, mark II.; if you consider it should be discarded from a grade eight test, mark it III. Make up your mind before turning the card over, and do not permit any previous opinion to alter yours. If you have any suggestion or criticisms to offer, such as 'ambiguous,' 'half-true, half-false,' etc., please indicate."

When the opinions were all recorded, the average opinion was taken as a sound judgment, and as a result, some items were discarded and others were re-worded.

In order that the items might be graded in difficulty, they were given then to an unselected group of forty grade nine students, who the previous June had been given grade eight standing. The degree of difficulty of each question was decided by the number of students who failed to answer it correctly. One question only was

answered correctly by all; so it was discarded as being too easy. There were no questions which every one missed, the smallest number of students answering any single question correctly being two.

A last revision was then made, and some questions discarded because they overlapped or because they indicated the correct answer to some other question. The list was thus reduced to one hundred and fifty.

It was then divided into three sections, A, B and C. Sections B and C were arranged in order of increasing difficulty, determined by the number who had failed to answer it correctly in the try-out test. Section A being a matching-test, it was obviously impossible to so arrange it.

In an effort to discover, if possible, the reliability of the test, it was divided into halves, called Test I. and Test II., with questions arranged thus:

	Test I.	Test II.
Item	1	2
	4	3
	5	6

etc., in order that the tests should be of relatively equal difficulty.

Test A was then given to one hundred and fifty students who the previous June had satisfactorily finished grade eight. The group was representative and unselected; including children of professional men, business men, farmers, office workers, railway employees, labborers and children of foreign parentage; it also included children who in grade eight attended rural schools, city schools and town schools, graded and ungraded.

The results were recorded and the students graded according to rank. Four days later, Test B was given to the same students and the students graded once more according to rank. Finding then the difference in rank of each student from Test A to Test B, the coefficient of correlation was worked out by Spearman's formula.[1]

According to this, the coefficient of correlation was found to be .84+ . This representing the coefficient

---

[1] $\rho = 1 - \frac{6\Sigma(V_x - V_y)^2}{N(N^2 - 1)}$

when N = the number of cases.

Holzinger, Karl J. Statistical Methods for Students in Education; Ginn & Company, New York, 1928, pp.278-280.

of correlation for each of the two halves of the test, the Spearman-Brown prophecy formula[1] was used to determine the coefficient of correlation for the full test. It was found to be .91.

This completed the preparation of the short-answer test. To test its worth and to compare it with that of the essay-type test, it was necessary to administer it to as many cases as possible. Classes are continually being promoted and moving on to a new grade and some difficulty was experienced in finding a sufficient number of teachers who were able to spare the time to take part in such a study at the end of the school year, when, obviously, the test should be given.

Through the courtesy of the Department of Education, however, a paper composed of one-half of the short-answer test and one-half of the essay-type test was given to the grade eight students of Manitoba in June, 1932. Three hundred and eighty-five of the answer papers collected were used by the writer as Study No. I. of this experiment.

---

[1] $r_2 = \frac{Nr_1}{1+(n-1)r_1}$ Ruch, G. N. op. cit. p. 418.

In 1933 the short-answer test and the essay-type test were given in full to two hundred and seven students in the schools of Brandon. This was the basis of Study No. 2.

In 1934 the short-answer test and a new essay-type test, called for convenience a "planned-paragraph" test, were given to seventy students of Brandon schools. This constitutes the basis of Study No. 3.

A full report of the method of procedure, the findings and the conclusions drawn from these three studies, follows in the subsequent chapters of this thesis.

CHAPTER IV.

Correlation of Tests and the Significance of the Same.

As has been reported in the last chapter, the examination in Canadian History given by the Department of Education in Manitoba in June, 1932, to grade eight students, was in two parts.[1] Part I. was one-half of the objective test prepared as reported in chapter three, above; while Part II. was composed of essay-type questions. Each part had a total value of fifty marks and the passing mark for the examination was fifty marks.

Study No. 1.

Three hundred and eighty-five of the answer papers from among those submitted were used as a basis for the first part of this study. The results were as follows:

Part I. objective, showed a correlation coefficient of .92, by previous experiment, reported in chapter three, when tested for reliability by the correlation of chance halves. It was felt that Part II. contained too few questions to warrant correlating halves, but when correlated with Part I. by raw scores, with the Product

---

[1] See Appendix A, page i.

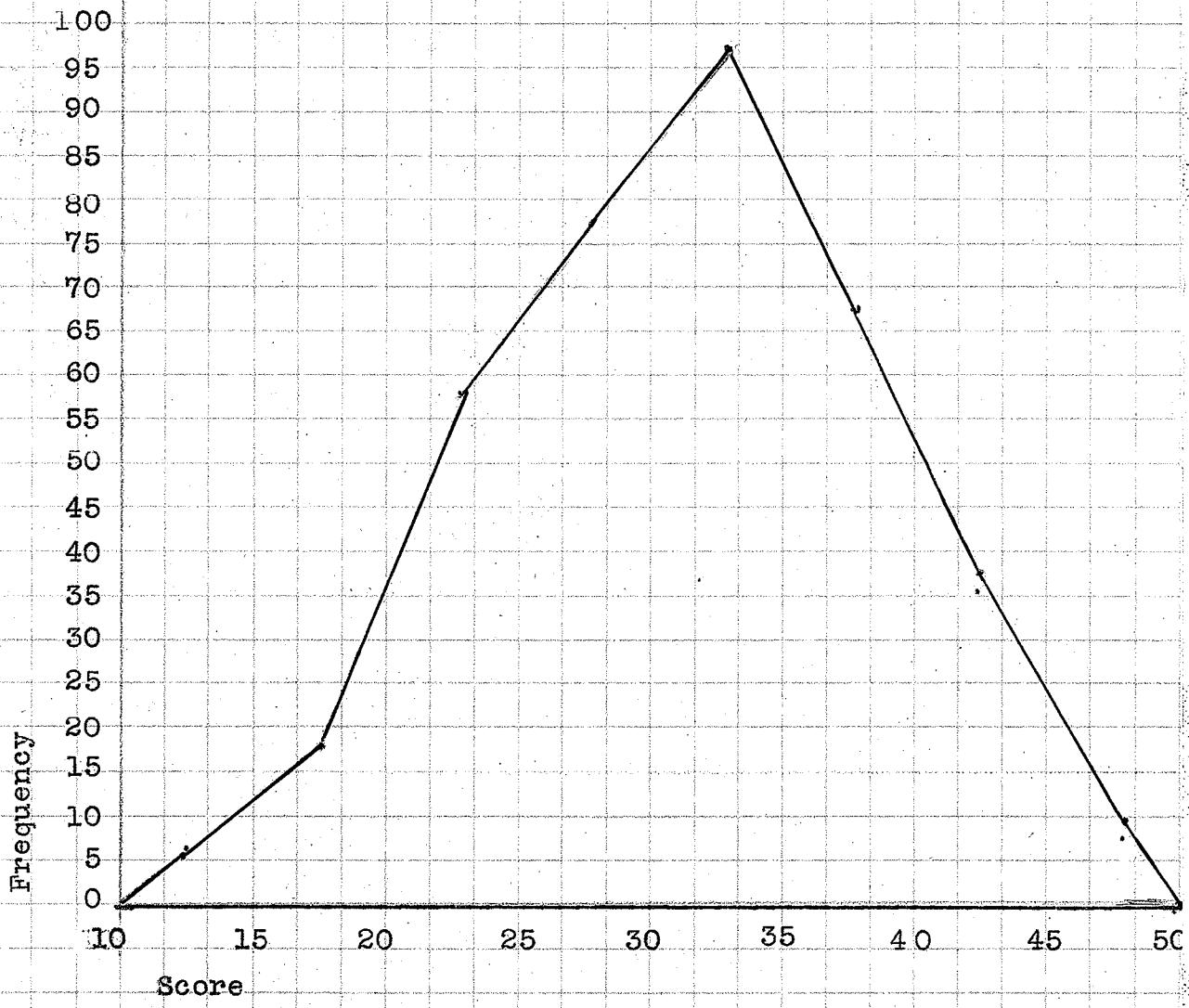


Figure 1.

Frequency Polygon of the Scores of 385 Students on Objective Test in Canadian History in June, 1932

Moment Formula,[1] the correlation coefficient of reliability showed to be .46, which would indicate the possibility of its reliability being low.

Other things of note which the study revealed, however, were (1) that whereas with papers that had been marked by two different markers, i.e.: one teacher had marked Part I. while another had marked Part II., the total scores approximated the normal curve; that among the papers marked by a single teacher the scores piled up very noticeably between 50 and 60; (2) that the low scores made on the objective questions were offset by the mark given on the essay-type questions among a great many of those papers receiving between 50 and 60 as a total mark; (3) that the frequency polygon of the scores on the objective part of the test indicates a leptokurtic curve with a slight negative skewness; but that the frequency polygon of the scores on the essay-type part of the test shows a badly broken curve, with a more marked negative skewness. (See Figures I. and II.)

---

[1] $$r = \frac{\Sigma fxyXY - \frac{(\Sigma fxX)(\Sigma fyY)}{N}}{\sqrt{\left\{\Sigma fxX^2 - \frac{(\Sigma fxX)^2}{N}\right\}\left\{\Sigma fyY^2 - \frac{(\Sigma fyY)^2}{N}\right\}}} = \frac{a}{\sqrt{bc}}$$

Holzinger, p. 151.

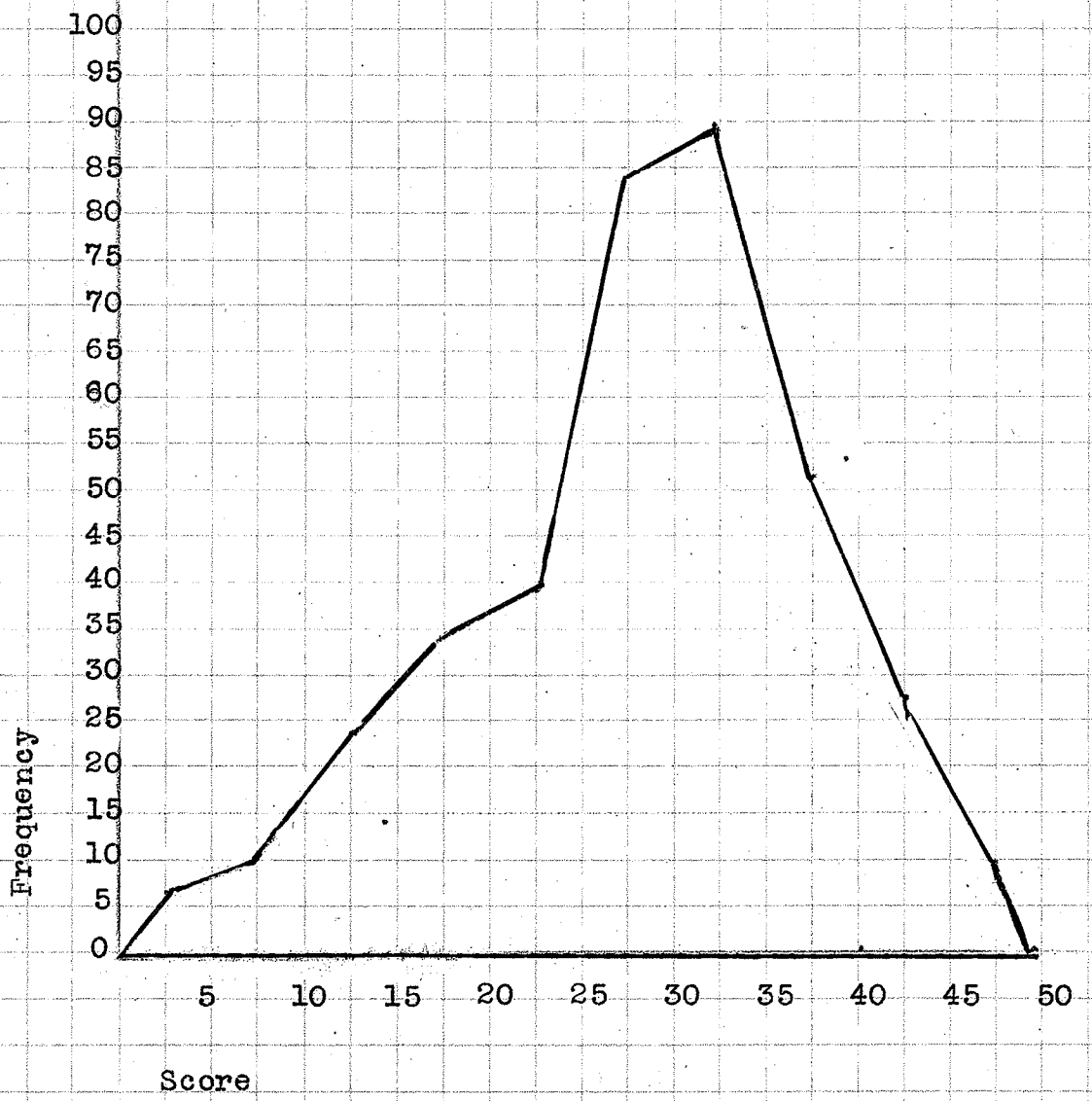


Figure 2.

Frequency Polygon of 385 Scores in Essay-Type Test in Canadian History, June, 1932

Study No. 2.

The short answer test given to Brandon students in September, 1933, was composed of the chance halves of the one-hundred and fifty questions prepared as described in Chapter III. The students had all been granted their Grade VIII. standing in June, 1933, therefore it was to be assumed that they would be able to meet the examination, if not as easily as in June, at least without great difficulty. They were told that they were going to have the examination and that they might prepare for it if they wished to do so.

Forty-five minutes were allowed for this test. The papers were then collected and marked, and the results recorded.

For the essay-type paper, given four days later, an hour and a half was allowed and again the papers were collected and marked by the same teachers and the results recorded. In the subsequent pages of this thesis, the short-answer test and the essay-type test will be referred to as Test A and Test B respectively (see Appendix A, pages x to xvi.).

Two weeks later the Otis Group Intelligence Test was given to the same classes, with the exception of one class where, owing to unavoidable circumstances, it was delayed. This reduced the number of cases to 157, but it was felt that the number was sufficiently large to be acceptable. The scores made by these 157 students were recorded.

Lastly, in order to obtain a composite rating of the 200 students, who took Tests A and B, based on the judgment of competent teachers, their teachers were called into conference in February. Since only one teacher had taught each class its Canadian History of grade eight, the teachers who are now teaching them history in grade nine were pressed into service, so that for each student there were three teachers to express a judgment. It was felt, generally speaking, that the student who masters the history of one grade also masters the history of another, and vice versa. The teachers were asked to rank the students according to their performance in class. Considerable difficulty was experianced in this connection. There was a marked tendency toward piling up the names at the upper end. Repeatedly it was found necessary to remind them that personality, or reward for effort, or language handicap were not factors to be considered; but that it was performance alone that was being judged. Since the students were

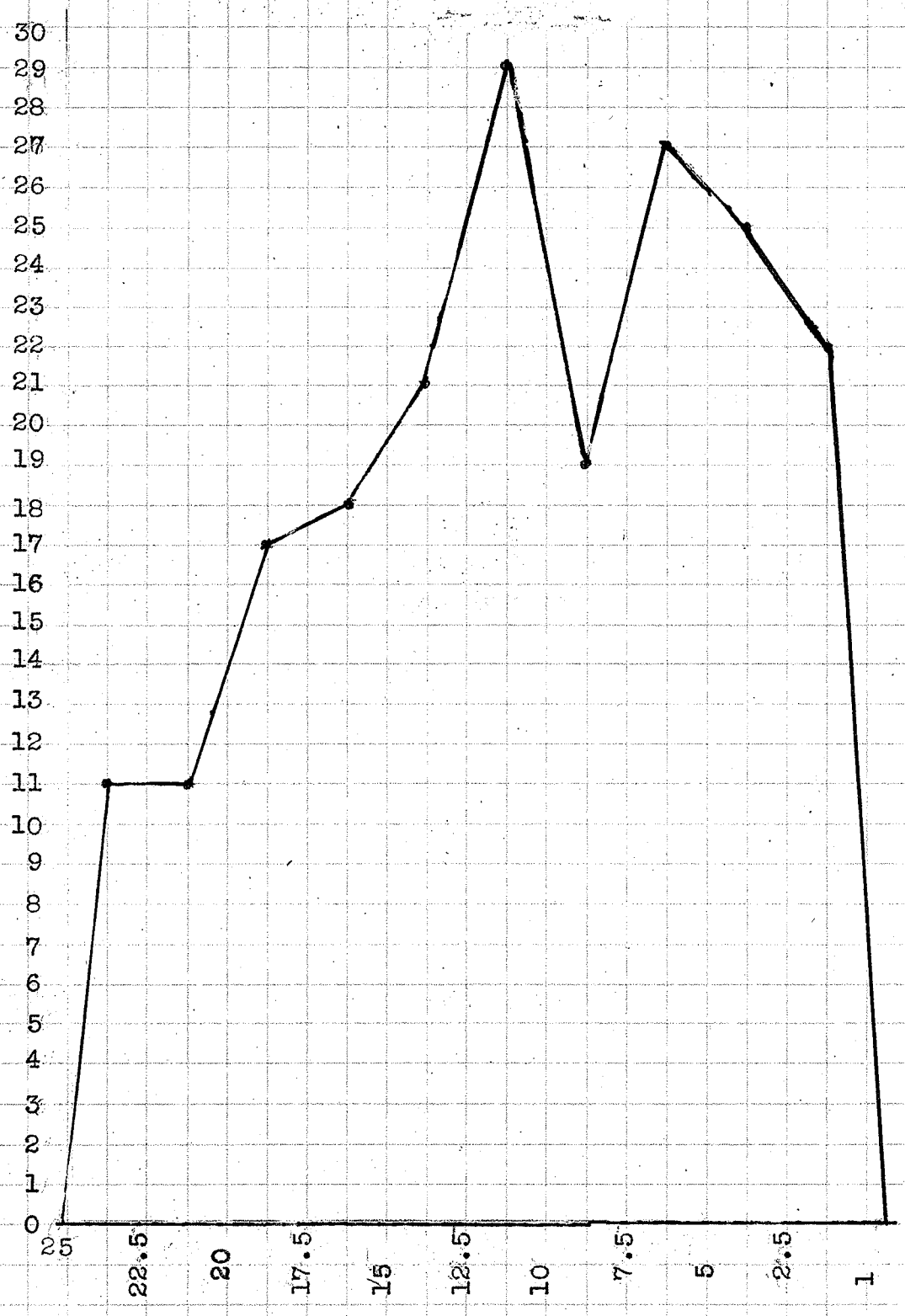


Figure 3.

Frequency Polygon of Teachers' Rating of
200 Pupils in Grade VIII Canadian History

not all in the same school, there was no means of ranking them as a whole group. To overcome this difficulty the suggestion was made that in each school they rank them as from 1 to 25, number 1 representing the best Canadian History student they had ever had, and number 25 the poorest. This plan worked satisfactorily, if we except the constant inclination to rank high all the better students. Five of the 200 were given the rank of 1, five others were given the rank of 25, but whereas 15 ranked as number 2, only 4 were ranked number 9. So that the curve was skewed negatively slightly, with a decided break just beyond the middle. (See Figure III.) Since these teachers have done practically no experimental work along these lines, it was felt that the judgment was as nearly accurate as it was possible to get from them. Ranking the five who were called number 1 as number 3; the succeeding fifteen as 13 (the middle score of the group); and so on, the 200 were ranked together. Similarly their scores on Test A and Test B were grouped and ranked in the same way, such rank bearing the number of its middle score. The difference in rank of each student was then determined as between Teachers' rating and Test A, and again as between Teachers' rating and Test B. Summing the squares

of those differences, computation was made of the coefficient of correlation according to Spearman's formulas[1] (1) and (2) given below.

With the scores and ranking thus obtained, certain computations have been made. The findings are herewith presented for consideration.

(1) Test A, being tested for reliability, by the ranking method on chance halves, showed a reliability coefficient of .92, when 167 cases were considered.

(2) Test B, tested on halves, also by the ranking method, in 200 cases, yielded α = .67, which gives r = .687.

(3) Part I. objective test showed a coefficient of correlation with Part II. essay-type test, of .46 when 385 cases were considered, of those who took the examination of the Department of Education in June, 1932. (See Table No. 4, Appendix C.)

(4)Test A, objective, showed a coefficient of correlation, on a distribution table, of .36 with Test B, essay-type, when 200 cases were considered.

---

[1] (1) $\alpha = 1 - \frac{6(\leq D^2)}{N(N^2-1)}$ Holzinger, Karl J. op. cit. Appendix A, p. 355.

(2) $r = 2 \sin (30^c \times \alpha)$ Otis, Arthur. Statistical Method in Education; World Book Co., Chicago, 1926; Table IV. Appendix II., p. 296.

(5) Test A, objective, being tested by means of a distribution table, for validity, against the scores made in 157 cases out of the 200 mentioned above, on the Otis Group Intelligence Test revealed a coefficient of correlation of .78, by means of Pearson's formula, given below.[1] Applying the formula for predicting the validity of a test of twice as many items, that coefficient becomes .81.[2]

(6) Test B, essay-type, being tested in the same way against the Otis scores of the same 157 cases, revealed a coefficient of .30, which the prediction formula raises to .328 on a test of twice the length.

The Standard Deviation of the scores made on the Objective Test is shown, when computed by formula,[3] to be 7.2, giving a probable error of estimated score to be 1.94

---

1 $$r = \frac{\Sigma fxyXY - \frac{(\Sigma fxX)(\Sigma fyY)}{N}}{\sqrt{\left[\Sigma fxX^2 - \frac{(\Sigma fxX)^2}{N}\right]\left[\Sigma fyY^2 - \frac{(\Sigma fyY)^2}{N}\right]}} = \frac{a}{\sqrt{bc}}.$$

Holzinger, Karl J., op.cit. p. 151.

2 $$r_{cn} = \frac{nr_{cz}}{\sqrt{n + n(n-1)r_{zz}}}.$$

3 $$S.D. = \left\{\sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2}\right\} h$$ Ibid. p. 348.

by formula.[1] (See Table 11, Appendix C.)

The Standard Deviation of the scores on the essay-type paper when computed by formula[2] is shown to be 21.10, with a probable error of estimated score of 10.38. (See Table 11, Appendix C.)

(7) Test A, measured against composite rating based on teachers' judgments of the 200 cases previously mentioned, by the ranking method gives a coefficient of .658 by formulas (1) and (2) on page 53, (footnote). This becomes, on a test of double the length, .671.

(8) Test B, measured against the same composite rating, with the same 200 cases, shows a correlation coefficient of .497.

Realizing that forgetfulness must be considered as a factor, when several months have elapsed between the study of the subject and the test, and especially when these months have been vacation months, the scores made on the Otis test were correlated on a distribution table with the scores made on an essay-type test five months earlier,

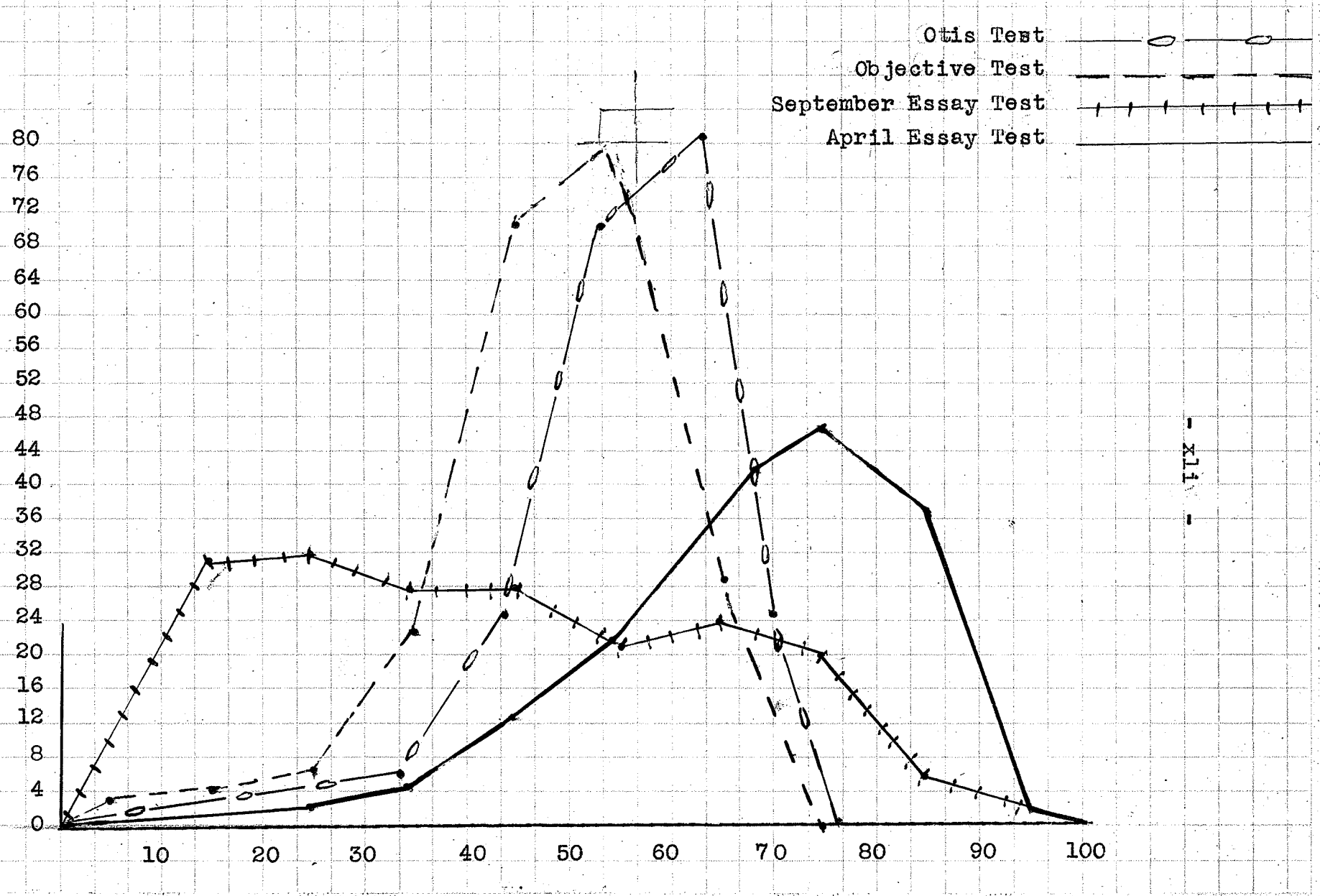
---

[1] $P.E._{est.} = .6745 \times S.D._{dist.} \times \sqrt{1 - r^2}$. Ibid. p. 350.

[2] $S.D. = \left\{ \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \right\} h.$ Ibid. p. 348.

Figure 4.

Frequency Polygons of Scores Made on Four Tests, Arranged for Comparison.



in the latter part of the spring term 1933, on the same Canadian History, by the same students. The coefficient of correlation with the scores on the Otis test was found to be even lower, .18. It is to be noted in this connection that whereas the polygon of the scores made in September on the essay-type test is such that the curve is positively skewed, that the polygon of the scores made on the same type of test on the same work five months earlier indicates a curve skewed negatively. (See Figure 4.)

To obtain a measure of consistency of rating by the essay-type test, these two sets of scores were correlated with each other on a distribution table. The coefficient was .49. Figure 4 shows the polygons of these two tests in contrast with the polygons of Test A and the Otis Test.

To verify the accuracy of the computations made, a colleague, to whom the statistical work is familiar, made independent calculations. The findings she obtained coincided with those presented here. The actual scores obtained in these tests are presented in Appendix.B.

CHAPTER V.

COMPARISONS BASED ON FORE-GOING RESULTS.

From the foregoing results, certain comparisons may be made and significant conclusions drawn.

Comparing the reliability of the objective and the essay-type tests, as shown by correlations on halves, we find that Test A, objective type, has a reliability of .92 against .68 of Test B, essay-type. When correlated against each other, the coefficient is .46, which would indicate that should our succeeding correlations show a high validity for Test A, they will at the same time indicate a very low validity for Test B. The chance halves correlation of Test B, being only .68, would support this argument, inasmuch as it indicates that its reliability is doubtful.

Again, taking as an authentic standard of validity the score of the Otis Intelligence Test, we find that the correlation of the objective test is .815 against a correlation of .32 with the essay-type test. It would therefore appear that the objective-type test has shown greater discriminative value for the recognition of intelligence.

Accepting as valid the composite judgment of a group of teachers competent to judge, it will be seen that in ranking the students by the two types of test, the objective test shows a ranking correlation of .67; the essay-type test, a ranking correlation of .49. It will be seen that the objective test is more in harmony with the judgments of teachers who are in a position to speak with some authority. As has been mentioned before, this composite judgment was not so satisfactory as it would be, were these teachers more accustomed to ranking their students in this manner, but the frequency polygons would indicate that the nearer that composite judgment approached the normal curve, the more nearly would it coincide with the objective-test results, and the higher would be the resulting correlation.

It has not been possible to make any computation of the consistency of rating when the objective test is used, inasmuch as the tests given to these students at the close of their term last spring (1933) was not objective. That the rating by the essay-type test in September was not consistent, however, with the rating at that time is evident from the dissimilarity of their

frequency polygons. (See Figure 4, page fifty-two..) Correlating the two scores, on a distribution table, shows a correlation coefficient of .48 between the scores made in April and those made in September by the same pupils, which would indicate that that type of test, at least, can not be recommended for the consistency of its rating.

Summarizing the conclusions, we may say:

(1) As an estimate of reliability, the objective test had a coefficient of correlation of .92; the essay-type test had a coefficient of correlation of .68. This indicates a greater degree of reliability in the objective test. The standard deviation from the average was 7.2 in the objective test, against 21.10 in the essay-type test; the probable error of the estimated score in the objective-type test was 1.94 as against 10.38 of a probable error in the essay-type test; that is to say, in a large distribution, the objective test may have more than five times the chance of scoring correctly than does the essay-type test.

The curve of distribution approximates a normal curve in the case of objective scores. That of essay-type scores is much skewed, in one case positively, in the other negatively. This would indicate that certain constant

factors were at work in one case to make the scores unduly great, and in the other, unduly small.

We may conclude that the objective test was more reliable.

(2) As a measure of validity: accepting as valid the composite judgment of competent teachers in this field, we find a correlation of .67 with this standard in the case of the objective test, and of .49 in the case of the essay-type test.

Accepting as a standard of validity the power to indicate the intelligence of the student, we find a correlation of .815 in the case of the objective test, and of .32 in the case of the essay-type test.

We may therefore conclude that the objective-type of test was more reliable than the essay-type; and, on the standards of teachers' ratings and discrimination of mental ability, more valid.

What general conclusions are we justified in drawing from the above? First, in respect to scope of the tests; Test A consisted of seventy-five questions as opposed to five questions on Test B. That is to say, fifteen times as often was the mind of the student probed, in Test A as in Test B, and "luck" was, by just that proportion, reduced.

Secondly, Test A has evidently dealt more justly with the pupils. The scores in the essay-type test, taken in September, were so low that not more than one-quarter of the pupils would have been given a pass on them. The teachers were asked to record the reaction of the pupils and they were unanimous in reporting, "Consternation." It was evident that the "notes" they had so carefully memorized a few months earlier, were completely gone from their memory, and they had, in most cases, nothing to put in the place of these notes. An inquiry sent to the teachers asking for their explanation of this condition, received the same answer from each. "The result of cramming for examinations; memorizing notes; and worrying over composition."

Test A, objective, did not produce the same results. Here the students were meeting facts about which apparently they felt reasonably sure. A vote taken from them on which of the two types of test they preferred, was overwhelmingly in favor of the objective test. It was more "fair," they said, by which, it would seem, they meant that they knew the facts and wished to be judged on them.

Thirdly, in respect to the time spent on tests. The building of the objective test required time and care; it was the product of real thought and study on the part of the test-maker, but the time saved in marking the papers more than repaid him. In this study the teachers who were marking the papers were asked to time themselves on each test and compare the length of time required to mark five essay questions with that required to mark seventy-five objective questions. They reported that the seventy-five objective questions were marked in approximately one-quarter of the time required to mark the five essay-questions. In one class, the papers were exchanged and the correct answers read aloud by the teacher. The whole set of approximately forty papers was marked (without an error) within fifteen minutes. That plan has the added advantage, if used during the term, of providing the pupils with a most effective review of their work. The fact that the test can be taken, papers exchanged and scored, in less time than the essay-type test can be written, gives the objective test a decided advantage.

Fourthly, in respect to the real worth of the score obtained: the scoring of the objective test was done in such a way that personal bias or opinion had no effect on

it; a clerk, a student or a teacher marking the paper would have given it the same score. The student was paid for exactly what he put on his paper. The validity of the objective test being higher than that of the essay-type test, it would indicate that the student was paid more in accordance with the extent of his knowledge of the subject on which he was being tested.

The essay-type test produces a subjective score, into which personal bias and opinion almost inevitably enter and which permits other factors, even less worthy, to enter. The fact that on the two halves of the departmental examination in June, 1932, the correlation was only .46, when in a large percentage of the cases considered, one marker had marked both halves, and when a below-50% mark meant a failure in the subject, points to a possible manipulating of marks, which is most undesirable. Apart from that consideration, the most sincere and earnest marker could not be sure whether he was marking History alone, or History plus Composition. The variance in the marks granted to a single paper by a number of markers working independently is too well known to be repeated here. When, too, the same students, writing on the same subject at two different times, less than five months apart, have an average drop of 49 marks on a 100-mark paper, as was the case in one class-

room where these tests were given, it makes the real worth of either score very doubtful.

Study No. 3.

The objection is raised, quite legitimately, that the short-answer test administered was carefully constructed, that a study was made of the most effective procedure to be followed in the building and validating of such a test, while the essay-type test administered represents the type of test set hurriedly and without any attempt at validation; and that a comparison of results is not justifiable. The point is granted; nevertheless, the fact remains that the essay-type test used was a fair sample of the average history test of the average history teacher in the schools of Manitoba to-day, as is indicated by the fact that it was compiled from the tests of the past fifteen years and was approved by a representative group of teachers engaged in that work and deemed competent teachers.

With the thought in mind, however, that the essay-type test might be made more effective as a measuring instrument if more thought and care were put into its construction, the writer made a third study, which, while less extensive than the former ones, seems to bear out that contention.

Each of five competent Grade VIII. History teachers was asked, in May, 1934, to construct the best essay-type question he could formulate on each of the six main topics dealt with in the history for the grade. The stipulation was made that the questions must provide an outline of the discussion, sufficient to give a student of the Grade VIII. age a clear idea of what was expected of him.

These questions, when submitted, were presented to a group of four teachers who have made a special study of testing and teaching history, for critical consideration, and from among the list of questions submitted a test-paper was compiled. (See Appendix A, Test C.) These questions provided the students with definite plans for their essay answers, and may be termed "planned-paragraph" questions.

Unfortunately, a group of only seventy students was available for the experiment. Nevertheless it was felt that even so small a group might indicate the trend of results.

Owing to the relatively small number of cases and the small number of questions, the test received no further test for reliability than that of the pooled judgment of competent people. For validity, however, it was tested against the scores made by the same students on the Otis Test and against the scores made by the same students on short-answer test A. The time allowed for the test was

one and a half hours, as had been allowed for the essay-type test in the former experiment. By the majority of the students the time was felt to be too short, although there had been no such expression of opinion in the case of the first essay-type test (Test B.).

The papers in this case were marked by three of the five who had marked the papers in study number 2; and the results were recorded. These results were as follows:

(1) When correlated with the scores made on the short-answer test, the coefficient of correlation was .60; P. E. .07.

(2) When correlated with the scores made on the Otis test, the coefficient of correlation was .62; P. E. .06.

Comparing these results with those obtained in study 2, it is found that (1) whereas the first essay-type test had a correlation of .46 with the short-answer test, the planned-paragraph test had a correlation of .60 and a standard deviation of 13.1 as contrasted with a standard deviation of 21.1 in the case of the first essay-type test; (2) the first essay-type test had a correlation of .30 with the Otis test; the planned-paragraph test had a correlation of .65, while the short-answer test had a correlation of .81 with the same test. Accepting the Otis Test as a standard of validity, it would appear that by providing

students, of the grade eight age at least, with a plan which will indicate just what phases of the topics their various paragraphs will develop, that the validity of the test is materially increased.

This experiment, limited though it was, would seem to indicate something that is of primary importance, namely: that by collaboration with others competent in the field, the essay-type question can be improved and made to elicit more definite answers, the value of which may be measured with greater accuracy. To achieve this, however, it would seem that teachers must first be made aware of the weakness of the essay-type test as we have been accustomed to see it, and of the possibility of constructing a better measuring instrument through thoughtful study and cooperation with others engaged in the same work.

Nor is it to be assumed that a short-answer test, as such, is invariably satisfactory. Without a thorough knowledge of the method of constructing it, and the will and patience to construct it carefully, the short-answer test-maker may produce something worse than useless. If they are to be used, a course in the construction of short-answer tests must be a part of teacher-training, otherwise only standardized tests should be used.

CHAPTER VI.

Summary and Conclusions.

According to the methods employed in securing and analyzing the data for this study, the writer feels warranted in drawing the following conclusions:

(1) A carefully prepared objective test gave a more valid estimate of pupil knowledge in Grade VIII. History than did the essay-type test used for the purpose of examining the same subject. This may be interpreted as reflecting on the quality of the essay-type questions used. However, it has been shown already that the questions selected for this paper were typical of those used for the Grade VIII. final examinations in general, during the period 1920-1932. Thus we are forced to arrive at either of two conclusions; e. g., (1) that the objective test is superior to the essay-type for the purpose of estimating pupil-knowledge in history, or (2) that the essay-type examination used during the period in question was not of the best quality.

(2) Since similar attention given to the preparation of the essay-type examination brought the pupil results for each type into closer relation, it would seem apparent that the wide difference in results obtained from Tests A and B was due more to the preparation and possibly to the

marking, of each test, rather than to the inherent superiority of one type over the other. In this connection it must be remembered that the opportunity to use the specially prepared essay-type examination was more limited than that afforded by Tests A and B, and therefore any conclusions drawn would not be exactly comparable. Sufficient was done, however, to indicate that the essay-type test may be strengthened significantly through limiting the range of questions and providing a clearer definition of their scope.

The findings arrived at from the second phase of this study would cause one to question the suitability of the essay-type examination papers used for Grade VIII. promotion during recent years, and would strengthen the opinion expressed by Dr. Tryon (page thirty-two) that"the present status of the written examination in the social studies would seem to indicate that improvement rather than abolishment is what is needed." As previously stated, Dr. Krey, after making a comparison of the respective values of the objective and essay-type tests, states, "For these reasons, the members of our department who use the objective-type test do not employ it for more than one-half the examination." As this procedure has been adopted in part in Manitoba, it might well be made the practice in general.

The results obtained by the one type of test might be used to advantage to serve as a check on the other. This would be especially effective in the marking of history papers.

In conclusion, let us consider very briefly the effect produced on history teaching throughout any community, by the type of examination given at the end of the term. So long as the success or failure of a student's year is to be reckoned by the score made on a written examination of some kind, that examination will be, very largely, the guide to the teacher. It would seem to be imperative, then, that it should be used to guide him wisely. Laborious memorizing of history notes is, apparently, an actual waste of time if the understanding gained cannot be expressed satisfactorily five months later. Again, if the questions on which the student is to write are inevitably to be confined to ten or twelve, indefinite in their scope, the tendency is certain to be to prepare for these ten or twelve. If the questions are numerous, and the answers are either right or wrong, the student must be trained to be definite in his statements and certain of his facts; rhetorical excellence will not cover a deficiency in history, nor will efficiency in

history be undervalued because of faulty composition. Teachers will have a guide to help them weigh and judge the ability of the students.

On the other hand, as Dr. Krey says, "beyond the reach of the objective test is the very quality which by title the objective test seems to scorn, that is, the subjective values." The objective type through over-emphasis on encyclopaedic knowledge tends to minimize the importance of training in the organization of ideas and takes from history teaching one of the reasons for its being on the curriculum. Due emphasis on factual information as the raw data upon which reasoning depends, combined with emphasis on the organization of that factual data to the point of a larger social understanding, should be the aim of history teaching; and the examinations set by those in authority should be such as will reveal whether that aim is being realized. By so doing, the examinations may become an important factor in the reform of methods of teaching history.

When, to-day, the school is being called upon to defend its right to the liberal support of an over-burdened people, every effort must be made to prove the

worth of what the school offers, and every effort made to teach and test, in the most effective way, whatever is worth teaching. Society is justified in looking to the school to lead the way out of the social morass into which the world has strayed. The school has the youth of the world in its hands and in that youth is the hope of the future; one might be tempted to say, the only hope. Teachers must teach worthily, must test accurately and must develop the ability to judge performance correctly. If, therefore, this study may have served, in any small measure, to point the way toward improved methods of testing one of the most valuable of all our subjects, one of the Social Studies, it will have been justified. It is in the hope that such may be the case that this thesis is very humbly submitted.

BIBLIOGRAPHY

Ballard, P. B. The New Examiner. Hodder and Stoughton, London, England, 1929.

Holzinger, Karl J. Statistical Methods for Students in Education. Ginn and Company, Chicago, 1928.

Odell, C. W. Educational Measurements in High School. The Century Company, N. Y., 1930.

Orleans and Seely. Objective Tests. World Book Company, Yonkers, N. Y., 1928.

Otis, Arthur S. Statistical Method in Educational Measurement. World Book Company, Yonkers, N. Y.,1926.

Ruch, G. M. The Objective or New-type Examination. Scott, Foresham and Company, Chicago, Ill., 1929.

Ruch, G. M. Improvement on the Written Examination. Scott, Foresham and Company, Chicago, Ill., 1924.

Ruch, G. M. and Others. Examination Methods in the Social Studies. Scott, Foresham and Company, Chicago, Ill., 1926.

Ruch and Stothard. Tests in High School Instruction. World Book Company, New York, N. Y., 1927.

Wood, B. D. Measurement in Higher Education. World Book Company, NewYork, N. Y., 1923.

Publications.

Educational Research Bulletin, Los Angeles, Vol. VII., 1927;

Historical Outlook, Vol. XVIII., XIX., XX.

School Review, Vol. XIX., XX., XXI., XXVI.

Year Books, 22nd and 26th, Public School Publishing Company, Bloomington, Ill.

APPENDIX A.

Department of Education Examination in
Canadian History, Grade Eight, June, 1932.

Candidate's Name .....
Address .....
Inspector's Name .....

DEPARTMENT OF EDUCATION
MANITOBA

EXAMINATIONS, JUNE,
1932

HISTORY
Tuesday, June 21st, 9.00 to 11.00

PART I

One hour is allowed for this portion of the paper. Promptly at ten o'clock the candidate will hand in Part I of the paper to the presiding examiner.

Section A

1. The column at the left below contains the correct dates of important events in Canadian History. The column at the right contains the names of ten important events. Read the names of the events carefully, then read the list of dates. Find the correct date for each event and place the number (not the date) in front of the name of the event. The first one is done correctly.

Date	Answer	Event
1. 1763	4	Christopher Columbus discovers America.
2. 1670	.....	Selkirk Settlers Come to Red River.
3. 1829	.....	Lord Durham Makes His Report.
4. 1492	.....	U. E. Loyalists Come to Canada.
5. 1858	.....	Canadian Pacific Railway Finished.
6. 1812	.....	Confederation of Canada.

7. 1885 ..... Welland Canal Opens Great Lakes to Shipping.

8. 1867 ..... Hudson's Bay Company Founded.

9. 1821 ..... Marquis Wheat Discovered.

10. 1903 ..... Alexander McKenzie Reaches Pacific.

11. 1793 ..... Canada Becomes a British Colony.

12. 1854 ....

13. 1839 ....

14. 1783 ....

2. The column at the left below gives the names of important men. The column at the right gives the names of ten events connected with the names of ten of the men. Read the events over and put in the answer column the number (not the name) of the man connected with that event. The first one is done correctly.

Men	Answer	Event
1. Miles Macdonnell	4	Discovery of America
2. Egerton Ryerson	....	Public Schools Introduced into Upper Canada.
3. John Cabot	....	Selkirk Settlement on Red River.
4. Christopher Columbus	....	Indian Torture of Prisoners.
5. Sir Guy Carleton	....	Beginning of Cabinet Government in Canada.
6. Père Lalemant	....	Rebellion of 1837.
7. General Murray	....	Responsible Government Established in Canada.

	Men	Answer	Event
8.	Captain Cook	....	Discovery of Newfoundland.
9.	Lord Elgin	....	Northwestern Exploration.
10.	Samuel Hearne	....	Building of the C. P. R.
11.	Champlain	....	Coming of the U. E. Loyalists.
12.	Lord Sydenham	....	.....
13.	Louis J. Papineau	....	.....
14.	Lord Strathcona	....	.....

Section B

3. There are 56 incomplete statements below. Five words or phrases are given after each statement. One of these five words, or phrases, makes the statement true. The other four are incorrect.

Read each incomplete statement carefully; decide which of the five possible words or phrases makes the truest sentence, then underline the word or phrase and write the number (not the word or phrase) in the parenthesis at the right. The first one is answered correctly.

1. America was discovered in 1492 by (1) Marco Polo, (2) Christopher Columbus, (3) Jacques Cartier, (4) Champlain, (5) Erickson .....(2)

2. The most warlike tribe of Indians in North America was the (1) Huron, (2) Cree, (3) Blackfeet, (4) Iroquois, (5) Algonquins .....( )

3. The French fur-traders who went off and lived in the woods were called (1) coureurs de bois, (2) habitants, (3) voyageurs, (4) half-breeds, (5) Metis .....( )

4. To meet the problem of lawlessness in the West, Canada created, in 1873, (1) The Strathcona Horse, (2) a standing army, (3) the Royal North-West Mounted Police, (4) organized law courts, (5) a jury .....( )

5. The real rush of settlers to British Columbia was the result of (1) the fur-trade, (2) the discovery of gold, (3) the fine climate, (4) the coming of the railway, (5) the fishing ( )

6. The first bishop of Quebec was (1) Laval, (2) Talon, (3) Daniels, (4) Roberval, (5) Frontenac ..... ( )

7. Manitoba's sea-port to-day is (1) Fort William, (2) Port Arthur, (3) Winnipeg, (4) Churchill, (5) Montreal ..... ( )

8. The man, among the French, who could best handle the Indians was (1) Frontenac, (2) Champlain, (3) La Salle, (4) Laval, (5) Talon ..... ( )

9. In 1534, Canada was claimed for France by (1) Jacques Cartier, (2) Christopher Columbus, (3) Henry Hudson, (4) Samuel Champlain, (5) Verendrye ..... ( )

10. British Columbia was first linked up with Canada by the completion of the (1) C. N. R., (2) C. P. R., (3) steamship lines, (4) Panama Canal, (5) telegraph system ..... ( )

11. Canada lost her great general, Sir Isaac Brock, in the battle of (1) Michilimackinak, (2) Detroit, (3) Queenston's Heights, (4) Moraviantown, (5) Lundy's Lane ..... ( )

12. The leader of the insurrection of the Metis was (1) Louis Riel, (2) Chief Crowfoot, (3) Cuthbert Grant, (4) William McDougall, (5) Poundmaker ..... ( )

13. The worst vice of the early settlers was (1) theft, (2) drunkenness, (3) greed, (4) dishonesty, (5) laziness ..... ( )

14. The colony at Louisiana was the result of the explorations of (1) Verendrye, (2) Groseilliers, (3) Radisson, (4) Alexander McKenzie, (5) La Salle ..... ( )

15. Laura Secord was the heroine of the battle of (1) Stoney Creek, (2) Beaver's Dams, (3) Lecolle Mills, (4) Chrystler's Farm, (5) Queenston's Heights ..... ( )

16. The Indian's dishes, Champlain said, were made of (1) porcelain, (2) deerskin, (3) birchbark, (4) bronze, (5) tin ..... ( )

17. The Boston Teaparty, was one of the causes of the (1) American Revolution, (2) French Revolution, (3) War of 1812, (4) Rebellion of 1837, (5) American Civil War ..... ( )

18. The prairies were first explored by (1) McKenzie, (2) Fraser, (3) Verendrye,(4)Selkirk, (5) La Salle ..... ( )

19. The Reciprocity Treaty of 1854 gave free trade on natural products between Canada and (1) England, (2) France, (3) U. S. A., (4) British Columbia, (5) Newfoundland ..... ( )

20. A great Indian Chieftain who helped Canada during the war of 1812 was (1) Pontiac, (2) Tecumseh, (3) Joseph Brant, (4) Cuthbert Grant, (5) White Eagle ..... ( )

21. The amazing growth of Canada between 1860 and 1880 was due to the introduction of (1) railways, (2) newspapers, (3) new trade regulations, (4) increased population, (5) established government ..... ( )

22. The first settlement on the prairies was made at (1) Regina, (2) Brandon, (3) Portage la Prairie, (4) Winnipeg, (5) Edmonton ..... ( )

23. The leader of the Reformers in Upper Canada in 1837 was (1) Louis J. Papineau, (2) William Lyon McKenzie, (3) Joseph Howe, (4) Bishop Strachan, (5) John B. Robinson .....( )

24. The American colonies rebelled because they objected to Britain's method of (1) raising an army, (2) government, (3) trading, (4) taxation, (5) colonization, for her colonies ..... ( )

25. The two most important Indian villages which Cartier visited were Hochelaga and (1) Donnaconna, (2) Sioux Falls, (3) Michilimackinak, (4) St. Croix, (5) Stadacona .....( )

26. The man who did most for the settlement of the Canadian West was (1) John A. Macdonald, (2) Samuel Hearne, (3) William McDougall, (4) Verendrye, (5) Selkirk ..... ( )

27. A Union Government was formed in Canada in 1917 by (1) Robert Baldwin, (2) Sir Robert Borden, (3) Sir John A. Macdonald, (4) Sir Wilfred Laurier, (5) Lord Durham ..... ( )

28. The boundary line between British Columbia and U. S. A. was settled in 1846 by (1) Washington Treaty, (2) Oregon Treaty, (3) Rush-Bagot Treaty, (4) Treaty of Utrecht, (5) Treaty of Paris ..... ( )

29. To please the new U. E. Loyalist settlers Britain in 1791, passed the (1) Quebec Act, (2) Constitutional Act, (3) B. N. A. Act, (4) Union Act, (5) Education Act .....( )

30. The U. E. Loyalists began the province of (1) Nova Scotia, (2) Quebec, (3) Ontario, (4) Manitoba, (5) British Columbia ..... ( )

31. In 1774, the British Parliament, in an effort to relieve discontent in Canada, passed the (1) Quebec Act, (2) Union Act, (3) B. N. A. Act, (4) Stamp Act, (5) Constitutional Act .. ( )

32. The "Right of Search" at 1812 meant the right to search vessels belonging to another nation for (1) runaway sailors, (2) liquor, (3) slaves, (4) treasure, (5) spies ..... ( )

33. The old town of York is to-day the city of (1) Kingston, (2) Toronto, (3) Niagara, (4) Halifax, (5) St. John ..... ( )

34. The first governor of Upper Canada was (1) General Murray, (2) Lord Dorchester, (3) Sir John Graves Simcoe, (4) William Lyon McKenzie, (5) Sir Francis Bond-head ..... ( )

35. The first conference which discussed Confederation met at (1) Toronto, (2) Montreal, (3) Quebec, (4) Charlettetown, (5) Halifax ..... ( )

36. The great road called Dundas Street, and the road called Yonge Street, in Toronto, were begun by (1) John Graves Simcoe, (2) Sir Guy Carleton, (3) Sir John A. Macdonald, (4) Frontenac, (5) The Indians ..... ( )

37. The "Order of Good Cheer" was organized in Acadia for (1) defence, (2) religious, (3) social, (4) political, (5) charitable, purposes ..... ( )

38. The result of equal representation in Parliament for Upper and Lower Canada, after the Union Act was passed was (1) rebellion, (2) agreement, (3) deadlock, (4) prosperity, (5) free trade .....( )

39. The leader of the Americans in the Revolutionary War was (1) Abraham Lincoln, (2) Theodore Roosevelt, (3) George Washington, (4) Lord North, (5) Sir Guy Carleton .....( )

40. The first party of settlers to Manitoba came by way of (1) The Great Lakes, (2) Hudson's Bay, (3) U. S. A., (4) Montreal, (5) Vancouver .....( )

41. Dollard des Ormeaux showed the Iroquois that France would not (1) settle, (2) abandon, (3) sell, (4) govern, (5) conquer, Canada .....( )

42. Port Royal was taken from the French in 1690 by (1) Sir Guy Carleton, (2) George Washington, (3) Sir William Phips, (4) General Wolfe, (5) William Pitt .....( )

43. The leader of the French in the struggle for supremacy in Hudson's Bay was (1) Radisson, (2) La Salle, (3) Montcalm, (4) D'Iberville, (5) Groseilliers ..... ( )

44. The centre of Iroquois attacks in Canada was around (1) Quebec, (2) Winnipeg, (3) Port Royal, (4) Montreal, (5) Hudson's Bay ..... ( )

45. A famous Indian Chieftain in Upper Canada when Simcoe came there was (1) Pontiac, (2) Tecumseh, (3) Joseph Brant, (4) White Eagle, (5) Poundmaker ..... ( )

46. Montreal was founded in 1642 by (1) Dollard, (2) Verendrye, (3) Champlain, (4) Maisonneuve, (5) Cartier ..... ( )

47. The Hudson's Bay Co., was formed on the advice of (1) Lord Strathcona, (2) Frontenac, (3) Lord Selkirk, (4) Henry Hudson, (5) Pierre Radisson ..... ( )

48. The Family Compact wanted Canada to be (1) like England, (2) independent, (3) Protestant, (4) open to all religions, (5) closed to the French ..... ( )

49. The "Trent" was a famous (1) battle, (2) newspaper, (3) fort, (4) Act of Parliament, (5) ship ..... ( )

50. The Union of the Maritime Provinces and Canada was strongly opposed by (1) Joseph Howe, (2) Sir John A. Macdonald, (3) Georges Cartier, (4) George Brown, (5) Charles Tupper ( )

51. The first European to reach the Pacific Ocean was (1) Balboa, (2) Drake, (3) Cook, (4) Raleigh, (5) Columbus ..... ( )

52. Montreal was founded for the purpose of (1) treating with, (2) Christianizing, (3) terrorizing, (4) making peace with, (5) defying, the Indians ..... ( )

53. The "Emperor of the Columbia" was (1) Wm. McDougall, (2) John McLaughlin, (3) George Vancouver, (4) James Douglas, (5) David Thompson ..... ( )

54. The first governor of the prairie provinces was (1) Lord Selkirk, (2) Sir George Simpson, (3) Miles MacDonnell, (4) Robert Semple, (5) Wm. McDougall ..... ( )

55. Sir John A. Macdonald thought a "protective tariff" would encourage Canadian (1) agriculture, (2) manufacturing, (3) literature, (4) mining, (5) trade with U. S. A. ..... ( )

56. The Dominion of Canada was formed by the (1) Union Act, (2) Confederation Act, (3) British North America Act, (4) Constitutional Act, (5) Peace of Paris ..... ( )

PART II.

(10 to 11 o'clock)

Candidates will detach this portion of the paper and will obtain booklets from the presiding examiner and answer the question paper in the usual way.

Stories of Canadian Leaders

12 x 2 = 24

4. Tell the story of any two of the following:

(a) Lord Selkirk.
(b) Frontenac.
(c) Lord Durham.
(d) Joseph Howe.

Stories of Canadian Life and Progress

13 x 2 = 26

5. Tell the story of any two of the following:

(a) Progress in agriculture in Manitoba since 1870.
(b) Progress in transportation in Canada since 1867.
(c) The terms of the British North American Act.
(d) The United Empire Loyalists in Upper Canada.

TESTS IN GRADE EIGHT CANADIAN HISTORY GIVEN TO THE STUDENTS OF GRADE NINE IN THE CITY OF BRANDON, IN SEPTEMBER, 1933

Part A. TEST A (OBJECTIVE)

1. The column at the left below contains the correct dates of important events in Canadian History. Read the events given in the column at the right carefully, then read the list of dates. Find the correct date for each event and place the number (not the date) in front of the event. The first one is done correctly.

	Date	Answer	Event
1.	1763	4	Christopher Columbus discovers America.
2.	1670	.....	Selkirk Settlers come to Red River.
3.	1829	.....	Canada becomes a British Colony.
4.	1492	.....	C. P. R. finished.
5.	1783	.....	Welland canal opened.
6.	1812	.....	U. E. Loyalists come to Canada.
7.	1885	.....	.....

2. The column at the left below contains the names of important men. The column at the right gives the names of events connected with five of the men. Read the events over, and put in the answer column the number (not the name) of the man connected with that event. The first one is done correctly.

Men	Answer	Events
1. Miles Macdonnell	(4)	Discovery of America.
2. Egerton Ryerson	(.)	Exploration of North West.

	Men	Answer	Events
3.	Pere Lalemant	...	Public Schools in Upper Canada.
4.	Christopher Columbus	...	Selkirk Settlements on Red River.
5.	Samuel Hearne	...	Discovery of Newfoundland.
6.	John Cabot	...	Indian torture of prisoners.
7.	Lord Elgin.	...	.....

Part B.

3. There are 28 incomplete statements below. Five words or phrases are given after each statement. Only one of these five words or phrases makes the sentence true. The other four are incorrect. Read each incomplete statement carefully; decide which of the five possible answers is correct, then underline the word or phrase and write the number (not the word or phrase) in the parenthesis at the right. The first one is answered correctly.

1. America was discovered in 1492 by (1) Marco Polo, (2) Christopher Columbus, (3) Jacques Cartier, (4) Champlain, (5) Erickson .....(2)

2. To meet the problem of lawlessness in the West, Canada created, in 1873, (1) The Strathcona Horse, (2) a standing army, (3) the Royal North-West Mounted Police, (4) organized law courts, (5) a jury .....( )

3. The real rush of settlers to British Columbia was the result of (1) the fur-trade, (2) the discovery of gold, (3) the fine climate, (4) the coming of the railway, (5) the fishing..( )

4. The man, among the French, who could best handle the Indians was (1) Frontenac, (2) Champlain, (3) La Salle, (4) Laval, (5) Talon .....( )

5. In 1534, Canada was claimed for France by (1) Jacques Cartier, (2) Christopher Columbus (3) Henry Hudson, (4) Samuel Champlain, (5) Verendrye .....( )

6. The leader of the insurrection of the Metis was (1) Louis Riel, (2) Chief Crowfoot, (3) Cuthbert Grant, (4) William McDougall, (5) Poundmaker .....( )

7. The worst vice of the early settlers was (1) theft, (2) drunkenness, (3) greed, (4) dishonesty, (5) laziness .....( )

8. The Indian's dishes, Champlain said, were made of (1) porcelain, (2) deerskin, (3) birch-bark, (4) bronze, (5) tin .....( )

9. The Boston Teaparty was one of the causes of the (1) American Revolution, (2) French Revolution, (3) War of 1812, (4) Rebellion of 1837, (5) American Civil War .....( )

10. A great Indian chieftain who helped Canada during the war of 1812 was (1) Pontiac, (2) Tecumseh, (3) Joseph Brant, (4) Cuthbert Grant, (5) White Eagle .....( )

11. The amazing growth of Canada between 1860 and 1880 was due to the introduction of (1) railways, (2) newspapers, (3) new trade regulations, (4) increased population, (5) established government .....( )

12. The American colonies rebelled because they objected to Britain's method of (1) raising an army, (2) government, (3) trading, (4) taxation, (5) colonization, for her colonies .....( )

13. The two most important Indian villages which Cartier visited were Hochelaga and (1) Donnaconna, (2) Sioux Falls, (3) Michilimackinak, (4) St. Croix, (5) Stadacona .....( )

14. The boundary line between British Columbia and U. S. A. was settled in 1846 by (1) Washington Treaty, (2) Oregon Treaty, (3) Rush-Bagot Treaty, (4) Treaty of Utrecht, (5) Treaty of Paris .....( )

15. To please the new U. E. Loyalist settlers Britain in 1791, passed the (1) Quebec Act, (2) Constitutional Act, (3) B. N. A. Act, (4) Union Act, (5) Education Act .....( )

16. The "Right of Search" at 1812 meant the right to search vessels belonging to another nation for (1) runaway sailors, (2) liquor, (3) slaves, (4) treasure, (5) spies .....( )

17. The old town of York is to-day the city of (1) Kingston, (2) Toronto, (3) Niagara, (4) Halifax, (5) St. John .....( )

18. The great road called Dundas Street, and the road called Yonge Street, in Toronto, were begum by (1) John Graves Simcoe, (2) Sir Guy Carleton, (3) Sir John A. Macdonald, (4) Frontenac, (5) the Indians .....( )

19. The "Order of Good Cheer" was organized in Acadia for (1) defence, (2) religious, (3) social, (4) political, (5) charitable purposes .....( )

20. The first party of settlers to Manitoba came by way of (1) The Great Lakes, (2) Hudson's Bay, (3) U. S. A., (4) Montreal, (5) Vancouver .....( )

21. Dollard des Ormeaux showed the Iroquois that France would not (1) settle, (2) abandon, (3) sell, (4) govern, (5) conquer, Canada ...( )

22. The centre of Iroquois attacks in Canada was around (1) Quebec, (2) Winnipeg, (3) Port Royal, (4) Montreal, (5) Hudson's Bay .....( )

23. A famous Indian Chieftain in Upper Canada when Simcoe came there was (1) Pontiac, (2) Tecumseh, (3) Joseph Brant, (4) White Eagle, (5) Poundmaker .....( )

24. The Family Compact wanted Canada to be (1) like England, (2) independent, (3) Protestant, (4) open to all religions, (5) closed to the French .....( )

25. The "Trent" was a famous (1) battle, (2) newspaper, (3) fort, (4) Act of Parliament, (5) ship .....( )

26. Montreal was founded for the purpose of (1) treating with, (2) Christianizing, (3) terrorizing, (4) making peace with, (5) defying, the Indians .....( )

27. The "Emperor of the Columbia" was (1) Wm. McDougall, (2) John McLaughlin, (3) George Vancouver, (4) James Douglas, (5) David Thompson ( )

28. The Dominion of Canada was formed by the (1) Union Act, (2) Confederation Act, (3) British North America Act, (4) Constitutional Act, (5) Peace of Paris ..... ( )

Part C.

Below are 38 true-false statements. Some of them are true and some of them are false. Read each statement carefully.. If it is true, place a (+) sign in the parenthesis; if it is false, place a (0) in the parenthesis. The first one is done correctly. Do not Guess.

1. Joliet and Marquette explored the Mississippi.. (+)
2. The Hudson's Bay Co. and the NorthWest Co. united in 1820-21, under the name of the Hudson's Bay Company ..... ( )
3. Lord Selkirk, by buying up a large tract of land from the Hudson's Bay Co., was able to start a settlement on the Red River ..... ( )
4. Very few children went to school in Upper Canada before 1800 ..... ( )
5. British Columbia's desire to remain part of the British Empire depended, in 1870 on the building of a trans-continental railway.. ( )
6. There were cities in North America before the Europeans came .....( )
7. The French of Quebec are more interested in France to-day than in Canada ..... ( )
8. Champlain's rash promise to help the Indians was responsible for the friendship of the French and the Iroquois ..... ( )
9. Robert Semple succeeded in driving off the Metis at Seven Oaks ..... ( )
10. The great attraction of the Western coast, at first, was the fur-trade ..... ( )

11. Baldwin and Lafontaine reformed University Education in Upper Canada ..... ( )
12. The U. E. Loyalists wanted a government in Canada which would permit them to vote ..... ( )
13. Britain acknowledged the independence of the American colonies in 1783 ..... ( )
14. Durham's report marks the beginning of a new period in Canadian History ..... ( )
15. The French of Quebec are the oldest Canadians in Canada ..... ( )
16. The lack of firmness of character made the Loyalists poor settlers ..... ( )
17. The settlers in Canada, a century ago, were too busy to be discontented ..... ( )
18. The British found in 1795, that the Indians on the Pacific Coast were being cheated in the Fur-Trade ..... ( )
19. If fair minded representatives of the British Parliament had met representatives of the colonies in conference in 1774, America would probably be British to-day ..... ( )
20. The Indians of North America did not have wagons because they did not know the use of the wheel ..... ( )
21. Canada had the same currency system as England until after the Great War ..... ( )
22. The Americans drew up the Declaration of Indpendence in 1776 to signify that they were free from France ..... ( )
23. The "Company of New France" founded by Richilieu brought many settlers to Canada ... ( )
24. Talon began industrial progress in Canada .... ( )
25. The Indians taught the Loyalists how to make clothing out of deerskin ..... ( )
26. The French Habitants of Acadia were much more comfortably situated than those in Canada ... ( )
27. The War of 1812 was caused by England's unfair methods of taxing her colonies ..... ( )
28. A large part of the credit for exploration and colonization of Canada is due to the missionary priests ..... ( )
29. When King Charles II. gave the Hudson's Bay Co. their charter to the monopoly of the fur-trade, he did not give them the ownership of any land ..... ( )

30. By the Municipal Act of 1849 all local affairs were put into the hands of the Dominion Government .....( )
31. Acadia remained a French possession until after the fall of Quebec, 1759 .....( )
32. The fair treatment of the French habitants after 1763 did not prevent them joining forces with the revolting colonies in 1775....( )
33. When Manitoba entered Confederation, her northern boundary was fixed at Hudson's Bay...( )
34. One of the early settlements in Manitoba was made on the shores of Lake Winnipeg by Icelanders.....( )
35. Slaves were bought and sold in Upper Canada before 1800 .....( )
36. No wheat was shipped out of Manitoba before 1885 .....( )
37. Upper and Lower Canada were united by the Union Act in 1840 .....( )
38. Lord Durham's success with the French of Canada was due to the support of a strong army( )

TEST B
Essay- Type

Canadian History

Answer Question 6 and any other four questions.
Value: 20 marks for each question.

1. Tell what you know of the development and settlement of New France, with special reference to the work of Samuel Champlain.

2. Give the date, Provisions, and results of the Constitutional Act.

3. Write a full note on the U. E. Loyalists and their settlements in the British Colonies.

4. Give the story of the Selkirk Settlement on the Red River.

5. Outline the courses leading up to the Rebellion of 1837 in (a) Upper Canada (b) Lower Canada.

6. Write notes on any 5 of the following: Tecumseh; Talon; Sir John A. Macdonald; Lord Durham; Lord Strathcona; Louis Riel; Joseph Howe; Sir George Simpson.

Test C.

Canadian History - Grade VIII.

The student will answer any five questions.

I. Discuss the work of Laval in Canada with respect to religion, morals, education, Indians and fur-trade.

O R

Discuss the work of Talon in Canada with respect to the form of government of his time; his duties as Intendant and his leadership in the colony.

II. Give an account of the U. E. Loyalists under the following heads:
(1) Who they were.
(2) Reasons for leaving The United States.
(3) Means of migration to Canada.
(4) Where they settled in Canada.
(5) The problem that was created by their coming.

III. Write a note on the development of Responsible Government in Canada under the following heads:
(1) What Lord Durham urged.
(2) Three demands of the people.
(3) The opposition to self-government; mentioning four people who opposed it.
(4) The work of three governors who advocated self-government.
(5) Three defects of the Union Act.

IV. Discuss Confederation of the Canadian Provinces, giving:
(1) Causes of Confederation.
(2) Steps leading to Confederation.
(3) The work of four important men.
(4) Expansion of Confederation.

V. Trace the story of the Red River Colony, following this outline:
(1) Lord Selkirk's proposal to the British Government and his bargain with the Hudson's Bay Company.
(2) Three groups of settlers and their journey to the Red River.

(Continued)

V. (Continued)

(3) Miles McDonnell; the Pemmican War, and its results.

VI. Give a short account of the explorations and discoveries of <u>any three</u> of the following:

La Verendrye,	Alexander McKenzie,
Samuel Hearne,	David Thompson,
La Salle,	Simon Fraser.

TABLE OF SPECIFICATIONS FOR CONSTRUCTION OF OBJECTIVE TEST. BASED ON ANALYSIS OF GRADE EIGHT HISTORY TEXT BOOK FOR MANITOBA, 1931-32

No.	Topic	Key Letter		Percentage of Items
I.	Early Civilization	E		6%
II.	The French Colony	F		20%
	Exploration		4%	
	Missions		4%	
	Military and Political		8%	
	Social		4%	
III.	The Making of Canada	C		40%
	Military		4%	
	Exploration and Settlement		9%	
	Political		9%	
	Trade and Industry		9%	
	Social		9%	
IV.	Manitoba	M		20%
	The Hudson's Bay Co.		5%	
	Military		5%	
	Social		5%	
	Economic		5%	
V.	British Columbia	B.C.		14%
	Exploration and Discovery		3%	
	Political		5%	
	Economic		6%	

APPENDIX B.

TABLE I.

TABLE OF SCORES

385 CASES IN DEPARTMENTAL EXAMINATIONS IN MANITOBA, 1932.

Student	Score in Part A.	Score in Part B.	Student	Score in Part A.	Score in Part B.
1	43	33	20	25	25
2	16	34	21	19	26
3	38	34	22	27	36
4	36	31	23	38	38
5	40	35	24	35	31
6	37	20	25	28	32
7	34	44	26	31	38
8	23	29	27	48	46
9	46	35	28	32	35
10	39	34	29	36	34
11	27	24	30	34	40
12	46	38	31	28	38
13	37	46	32	30	32
14	23	33	33	32	36
15	39	31	34	30	40
16	39	39	35	26	38
17	29	34	36	31	34
18	41	16	37	22	33
19	41	37	38	42	34

TABLE OF SCORES - DEPARTMENTAL EXAMINATIONS
(Continued)

Student	Score in Part A.	Score in Part B.	Student	Score in Part A.	Score in Part B.
39	36	38	59	31	41
40	33	39	60	37	35
41	40	34	61	33	33
42	27	28	62	26	28
43	26	33	63	36	38
44	22	31	64	34	31
45	35	43	65	40	40
46	46	45	66	35	50
47	35	35	67	30	26
48	39	36	68	46	39
49	45	33	69	33	29
50	43	28	70	35	29
51	24	18	71	28	16
52	31	33	72	44	11
53	43	30	73	43	19
54	45	36	74	43	46
55	43	31	75	23	33
56	42	39	76	34	36
57	43	37	77	24	37
58	30	40	78	12	4

TABLE OF SCORES - DEPARTMENTAL EXAMINATIONS
(Continued)

Student	Score in Part A.	Score in Part B.	Student	Score in Part A.	Score in Part B.
79	28	34	99	32	21
80	26	19	100	21	18
81	28	28	101	28	42
82	14	9	102	31	37
83	18	15	103	35	40
84	23	27	104	41	35
85	19	33	105	20	30
86	17	11	106	41	34
87	25	35	107	27	23
88	15	2	108	37	29
89	38	36	109	28	25
90	32	32	110	30	29
91	43	39	111	47	43
92	31	29	112	22	30
93	38	30	113	27	34
94	25	30	114	32	29
95	34	30	115	44	41
96	33	37	116	31	39
97	33	31	117	26	30
98	29	35	118	30	34

TABLE OF SCORES - DEPARTMENTAL EXAMINATIONS
(Continued)

Student	Score in Part A.	Score in Part B.	Student	Score in Part A.	Score in Part B.
119	26	28	141	34	44
120	22	20	142	28	16
121	32	33	143	17	40
122	24	32	144	34	45
123	29	32	145	45	39
124	30	30	146	42	46
125	33	0	147	47	43
126	36	39	148	26	25
127	29	34	149	39	39
128	34	31	150	35	36
129	34	33	151	35	25
130	29	38	152	39	15
131	14	0	153	37	24
132	46	26	154	34	30
133	44	26	155	33	33
134	40	44	156	37	33
135	32	35	157	21	38
136	32	28	158	21	29
137	27	28	159	38	44
138	26	26	160	30	48
139	22	30	161	13	39
140	20	18	162	20	34

TABLE OF SCORES - DEPARTMENTAL EXAMINATIONS
(Continued)

Student	Score in Part A.	Score in Part B.	Student	Score in Part A.	Score in Part B.
163	22	38	181	43	46
164	16	14	182	25	25
165	30	25	183	31	23
163	22	38	184	42	41
164	16	14	185	35	38
165	30	25	186	36	29
166	22	4	187	38	26
167	40	26	188	39	30
168	45	28	189	45	33
169	31	33	190	27	26
170	37	25	191	26	24
171	34	19	192	39	33
172	47	39	193	34	33
173	35	29	194	33	30
174	38	30	195	36	28
175	32	11	196	37	25
176	28	22	197	41	38
177	29	21	198	36	33
178	39	46	199	28	24
179	31	29	200	21	18
180	32	22	201	36	34

TABLE OF SCORES - DEPARTMENTAL EXAMINATIONS
(Continued)

Student	Score in Part A.	Score in Part B.	Student	Score in Part A.	Score in Part B.
202	33	18	223	37	36
203	17	9	224	23	33
204	33	33	225	36	49
205	28	30	226	21	30
206	27	34	227	24	33
207	32	35	228	26	27
208	34	33	229	37	44
209	32	45	230	39	46
210	20	27	231	22	31
211	27	29	232	19	44
212	15	20	233	46	46
213	29	41	234	42	43
214	21	29	235	28	42
215	24	29	236	28	35
216	21	33	237	25	44
217	30	40	238	20	5
218	30	20	239	23	18
219	23	36	240	30	18
220	25	26	241	23	13
221	27	30	242	28	7
222	28	30	243	33	14

TABLE OF SCORES - DEPARTMENTAL EXAMINATIONS
(Continued)

Student	Score in Part A.	Score in Part B.	Student	Score in Part A.	Score in Part B.
244	25	19	266	30	23
245	31	8	267	29	34
246	26	14	268	36	24
247	27	26	269	39	25
248	29	10	270	27	32
249	23	10	271	29	25
250	22	12	272	33	28
251	28	14	273	35	23
252	30	18	274	35	25
253	37	26	275	37	37
254	28	9	276	35	21
255	23	11	277	35	33
256	22	18	278	33	34
257	37	19	279	21	27
258	32	15	280	38	37
259	36	23	281	29	21
260	31	31	282	31	20
261	33	19	283	37	34
262	17	14	284	36	24
263	33	37	285	27	31
264	32	26	286	33	31
265	39	32	287	28	29

TABLE OF SCORES - DEPARTMENTAL EXAMINATIONS
(Continued)

Student	Score in Part A.	Score in Part B.	Student	Score in Part A.	Score in Part B.
288	31	31	310	35	27
289	33	25	311	29	31
290	35	29	312	45	40
291	33	30	313	28	33
292	34	36	314	37	26
293	38	34	315	27	31
294	35	28	316	31	31
295	29	41	317	40	31
296	44	42	318	31	29
297	23	9	319	33	30
298	41	30	320	27	14
299	31	32	321	34	27
300	35	34	322	31	28
301	29	13	323	19	13
302	23	37	324	37	40
303	24	18	325	25	17
304	31	28	326	36	26
305	35	30	327	31	31
306	35	33	328	32	34
307	33	19	329	20	13
308	22	10	330	35	26
309	31	20	331	32	43

TABLE OF SCORES - DEPARTMENTAL EXAMINATIONS
(Continued)

Student	Score in Part A.	Score in Part B.	Student	Score in Part A.	Score in Part B.
332	41	41	353	23	29
333	31	32	354	22	21
334	30	26	355	29	27
335	27	26	356	39	39
336	46	33	357	26	26
337	39	24	358	25	11
338	21	17	359	25	14
339	32	41	360	21	27
340	27	26	361	23	24
341	33	29	362	19	24
342	42	37	363	19	18
343	36	36	364	31	23
344	42	37	365	26	20
345	15	17	366	18	22
346	41	36	367	31	20
347	33	32	368	35	26
348	35	26	369	24	22
349	33	45	370	23	29
350	31	25	371	18	26
351	40	36	372	17	12
352	25	13	373	20	13

TABLE OF SCORES - DEPARTMENTAL EXAMINATIONS
(Continued)

Student	Score in Part A.	Score in Part B.	Student	Score in Part A.	Score in Part B.
374	17	21	380	23	18
375	21	16	381	19	14
376	33	23	382	32	32
377	28	32	383	28	27
378	23	22	384	25	5
379	41	23	385	22	10

TABLE 2.

SCORES OF 200 PUPILS IN TESTS GIVEN IN BRANDON SCHOOLS, 1933

Case	Score in Objective Test, 200 cases	Score in Essay Test, 200 cases	Score in Otis Test, 157 cases
1	68	80	72
2	67	42	73
3	65	72	65
4	64	68	62
5	64	56	73
6	64	72	68
7	63	25	65
8	62	50	70

TABLE OF SCORES - BRANDON
TESTS (CONT'D)

Cases	Score in Objective Test, 200 cases	Score in Essay Test, 200 cases	Score in Otis test, 157 cases
9	62	37	69
10	62	73	61
11	62	65	68
12	61	74	63
13	61	48	..
14	61	33	66
15	60	75	62
16	60	42	53
17	60	35	54
18	60	62	72
19	60	68	66
20	59	38	64
21	59	46	68
22	59	46	71
23	59	23	64
24	59	52	55
25	58	32	69
26	58	55	66
27	58	12	68
28	58	44	66
29	58	38	56

TABLE OF SCORES - BRANDON TESTS (CONTINUED)

Case	Score in Objective Test, 200 cases	Score in Essay Test, 200 cases	Score in Otis test, 157 cases
30	58	60	69
31	57	53	67
32	57	66	60
33	57	39	62
34	57	78	66
35	56	28	63
36	56	32	67
37	56	27	55
38	56	6	69
39	56	18	66
40	56	68	66
41	56	60	50
42	56	35	64
43	56	70	..
44	56	70	61
45	56	72	71
46	55	63	67
47	55	44	66
48	55	72	66
49	55	70	57
50	55	55	59

TABLE OF SCORES - BRANDON TESTS (CONTINUED)

Case	Score in Objective Test, 200 cases	Score in Essay Test, 200 cases	Score in Otis Test, 157 cases
51	55	44	55
52	55	25	61
53	55	67	62
54	55	50	70
55	55	60	68
56	55	61	65
57	55	70	62
58	55	52	46
59	55	52	68
60	54	38	46
61	54	67	40
62	54	36	57
63	54	22	51
64	54	72	68
65	54	75	50
66	53	18	58
67	53	78	63
68	53	51	64
69	53	84	59
70	53	30	60
71	53	50	63

TABLE OF SCORES - BRANDON TESTS (CONTINUED)

Case	Scores in Objective Test, 200 cases	Scores in Essay Test, 200 cases	Score in Otis Test, 150 Cases
72	52	68	68
73	52	66	63
74	52	64	52
75	51	12	72
76	51	12	45
77	51	19	40
78	51	70	48
79	51	70	67
80	51	30	63
81	51	31	63
82	51	40	48
83	51	82	65
84	51	42	51
85	50	57	68
86	50	35	..
87	50	40	49
88	50	24	60
89	50	10	51
90	50	40	55
91	50	45	59
92	50	50	57

TABLE OF SCORES - BRANDON TESTS - (CONTINUED)

Cases	Scores in Objective Test, 200 cases	Scores in Essay Test, 200 cases	Scores in Otis Test, 157 cases
93	50	62	59
94	50	52	..
95	49	31	56
96	49	18	62
97	49	29	62
98	49	24	51
99	49	75	54
100	49	25	57
101	49	65	63
102	49	40	63
103	49	44	..
104	49	70	..
105	48	16	41
106	48	29	64
107	48	6	60
108	48	14	..
109	48	20	60
110	48	48	62
111	48	65	63
112	47	11	59

TABLE OF SCORES - BRANDON TESTS - (CONTINUED)

Case	Scores in Objective Test, 200 cases	Scores in Essay Test, 200 cases	Scores in Otis Test, 157 Cases
113	47	11	59
114	47	27	57
115	47	82	65
116	47	40	38
117	47	30	61
118	47	52	65
119	47	72	63
121	46	8	57
122	46	45	58
123	46	25	36
124	46	20	63
125	46	55	..
126	45	43	56
127	45	18	39
128	45	34	..
129	45	42	..
130	45	19	59
131	45	8	..
132	45	22	57
133	45	50	55
134	45	35	..
135	45	38	53

TABLE OF SCORES - BRANDON TESTS - (CONTINUED)

Case	Scores in Objective Test, 200 cases	Scores in Essay Test, 200 cases	Scores in Otis Test, 157 Cases
136	45	38	43
137	44	15	41
138	44	34	43
139	44	17	..
140	44	68	..
141	44	48	..
142	44	62	59
143	44	40	51
144	44	73	70
145	44	43	56
146	43	17	60
147	43	82	43
148	43	50	47
149	43	60	48
150	43	67	55
151	42	12	..
152	42	39	..
153	42	22	..
154	42	68	52
155	42	45	29

TABLE OF SCORES - BRANDON TESTS - (CONTINUED)

Case	Scores in Objective Test, 200 cases	Scores in Essay Test, 200 cases	Scores in Otis Test 157 cases
156	42	52	..
157	42	62	59
158	42	42	..
159	41	25	..
160	41	7	..
161	41	14	55
162	41	84	37
163	41	78	..
164	41	20	48
165	41	62	..
166	41	61	..
167	41	55	59
168	41	28	50
169	40	14	48
170	40	6	53
171	40	25	53
172	40	68	..
173	39	38	..
174	38	17	..
175	38	15	55
176	38	20	..
177	37	17	48

TABLE OF SCORES - BRANDON TESTS - (CONTINUED)

Case	Scores in Objective Test, 200 cases	Scores in Essay Test, 200 cases	Scores in Otis Test, 157 cases
178	37	33	..
179	37	31	..
180	37	21	..
181	37	20	43
182	36	32	52
183	35	16	42
184	35	31	..
185	35	40	..
186	34	50	50
187	34	50	..
188	33	35	..
189	33	28	54
190	32	40	..
191	32	19	61
192	32	5	..
193	30	62	..
194	29	20	..
195	29	5	..
196	28	38	30
197	28	3	42

TABLE OF SCORES - BRANDON TESTS - (CONCLUDED)

Case	Scores in Objective Test, 200 Cases	Scores in Essay Test, 200 cases	Scores in Otis Test, 157 Cases
198	22	60	41
199	17	13	..
200	8	45	,.

TABLE 3

SCORES IN GRADE VIII, 1934,
PARAGRAPH ESSAY

Student	Objective Test	Planned Paragraph	Otis Test
1	91	70	88
2	87	63	92
3	85	51	69
4	85	65	88
5	85	42	93
6	85	54	84
7	85	55	79
8	85	60	71
9	85	70	83
10	84	70	65
11	84	68	73
12	84	58	75
13	84	46	65
14	84	59	69
15	84	56	77
16	81	63	80
17	81	72	69
18	81	47	76
19	81	45	64
20	81	70	72
21	81	55	71
22	81	75	71
23	80	32	63
24	80	52	81
25	80	35	40
26	79	38	68
27	79	61	68
28	77	64	76
29	77	71	91
30	77	62	91
31	77	75	54
32	77	70	61
33	76	70	85
34	76	46	72
35	76	66	82
36	76	53	75
37	76	51	76
38	76	55	63

TABLE 3

(Continued)

Student	Objective Test	Planned Paragraph	Otis Test
39	76	50	67
40	76	33	53
41	76	47	67
42	75	55	77
43	73	50	81
44	73	18	64
45	72	34	81
46	72	70	88
47	72	33	65
48	71	43	65
49	71	34	64
50	71	54	61
51	71	39	69
52	69	66	91
53	69	38	58
54	69	49	76
55	69	48	63
56	68	20	54
57	65	36	68
58	63	68	83
59	63	63	95
60	61	43	72
61	61	60	83
62	61	53	60
63	60	32	72
64	60	65	81
65	57	18	57
66	56	47	79
67	51	50	61
68	49	15	60
69	49	35	64
70	43	16	49

APPENDIX C

TABLE I.

RANKING CORRELATION TABLE BASED ON HALVES OF ESSAY-TYPE TEST

Student	Rank in X Test	Rank in Y Test	Difference	D^2
1	1	3	-2	4
2	2	5	-3	9
3	3	22	-19	361
4	4	19	-15	225
5	5	2	+3	9
6	5	12	-7	49
7	7	1	6	36
8	7	4	3	9
9	9	44	-35	1225
10	9	38	-29	841
11	11	8	+3	9
12	11	9	+2	4
13	13	31	-18	324
14	14	45	-31	961
15	14	15	-1	1
16	14	26	-12	144
17	14	31	-17	289
18	14	12	+2	4
19	14	15	-1	1
20	20	28	-8	64
21	20	35	-15	225
22	20	45	-25	625
23	23	31	-8	64
24	23	81	-58	3364
25	23	10	+13	169
26	23	10	+13	169
27	23	52	-29	841
28	23	60	-37	1369
29	29	45	-16	256
30	29	19	+10	100
31	29	12	+17	289
32	29	22	+7	49
33	33	15	+18	324
34	33	28	+5	25
35	33	6	+27	729
36	33	36	-3	9
37	33	38	-5	25
38	33	76	-43	1849

TABLE I. (Continued)

Student	Rank in X Test	Rank in Y Test	Difference	D^2
39	39	76	-37	1369
40	40	28	+12	144
41	40	38	+2	4
42	42	64	-22	484
43	43	45	-2	4
44	43	45	-2	4
45	43	35	+8	64
46	43	31	+12	144
47	43	52	-9	81
48	48	52	-4	16
49	48	60	-12	144
50	48	64	-16	256
51	48	19	+29	841
52	48	72	-24	576
53	48	85	-37	1369
54	48	60	-12	144
55	48	52	-4	16
56	48	81	-33	1089
57	57	38	+19	361
58	57	76	-19	361
59	57	52	+5	25
60	57	64	-7	49
61	61	64	-3	9
62	61	52	+9	81
63	61	64	-3	9
64	64	7	+57	3249
65	64	28	+36	1296
66	64	88	-24	576
67	64	72	-8	64
68	64	64	0	0
69	69	52	+17	289
70	69	22	+45	2025
71	69	64	+5	25
72	68	81	-12	144
73	69	52	+17	289
74	69	15	+54	2916
75	75	72	+3	9
76	75	95	-20	400
77	75	27	+48	2304
78	75	72	+3	9
79	75	38	+37	1369

TABLE 2

RANKING CORRELATION BETWEEN TEACHERS' RATING AND SHORT-ANSWER TEST

Pupil No.	A	B	D	D^2
1	3	5	2	4
2	3	1	2	4
3	3	40	37	1369
4	3	52.5	49.5	2450.25
5	3	9.5	5.5	30.25
6	12	27.5	15.5	240.25
7	12	89.5	77.5	6006.25
8	12	52.5	40.5	1640.25
9	12	9.5	2.5	6.25
10	12	32.5	20.5	420.25
11	12	7	5	25
12	12	3	9	81
13	12	13	1	1
14	12	5	7	49
15	12	52.5	40.5	1640.25
16	12	79.5	67.5	4556.25
17	12	52.5	40.5	1640.25
18	12	2	10	100
19	22	52.5	30.5	930.25
20	22	9.5	12.5	156.25
21	22	5	17	289
22	22	62.5	40.5	1640.25
23	22	73	51	2601
24	22	27.5	5.5	30.25
25	22	40	18	324
26	31	22	9	81
27	31	9.5	21.5	462.25
28	31	40	9	81
29	31	27.5	3.5	12.25
30	31	141	110	12100
31	31	22	9	81
32	31	62.5	31.5	992.25
33	31	40	9	81
34	31	17	14	196
35	31	40	9	81
36	31	73	42	1764
37	41.5	32.5	9	81
38	41.5	73	31.5	992.25
39	41.5	52.5	11	121
40	41.5	116	74.5	5550.25
41	41.5	17	24.5	600.25

TABLE 2

(Continued)

Pupil No.	A	B	D	D^2
42	41.5	89.5	48	2304
43	41.5	52.5	11	121
44	41.5	73	31.5	992.25
45	41.5	17	24.5	600.25
46	41.5	163	121.5	14762.25
47	51	13	38	1444
48	51	79.5	28.5	812.25
49	51	32.5	18.5	342.25
50	51	68.5	17.5	306.25
51	51	123	72	5184
52	51	89.5	38.5	1482.25
53	51	116	65	4225.
54	51	108	57	3249
55	51	52.5	1.5	2.25
56	62	17	45	2025
57	62	73	11	121
58	62	89.5	27.5	756.25
59	62	89.5	27.5	756.25
60	62	52.5	9.5	90.25
61	62	52.5	9.5	90.25
62	62	27.5	34.5	1190.25
63	62	22	40	1600
64	62	89.5	27.5	756.25
65	62	52.5	9.5	90.25
66	62	40	22	484
67	62	141	79	6241
68	62	148	86	7396
69	73	62.5	10.5	110.25
70	73	32.5	40.5	1640.25
71	73	32.5	40.5	1640.25
72	73	131	58	3364
73	73	99.5	26.5	702.25
74	73	40	33	1089
75	73	79.5	6.5	42.25
76	73	52.5	20.5	420.25
77	73	52.5	20.5	420.25
78	79.5	108	28.5	812.25
79	79.5	99.5	20	400
80	79.5	163	83.5	6972.25
81	79.5	131	52.5	2756.25
82	87	32.5	54.5	2970.25

TABLE 2

(Continued)

Pupil No.	A	B	D	D^2
83	87	108	21	441
84	87	73	14	196
85	87	108	21	441
86	87	184	97	9409
87	87	131	44	1936
88	87	154.5	67.5	4556.25
89	87	154.5	67.5	4556.25
90	87	52.5	34.5	1190.25
91	87	68.5	18.5	342.25
92	87	170.5	83.5	6972.25
93	97.5	179	81.5	6642.25
94	97.5	123	25.5	650.25
95	97.5	40	57.5	3306.25
96	97.5	131	33.5	1122.25
97	97.5	131	33.5	1122.25
98	97.5	141	43.5	1892.25
99	97.5	141	43.5	1892.25
100	97.5	148	50.5	2550.25
101	97.5	131	33.5	1122.25
102	97.5	193	95.5	9120.25
103	108.5	108	.5	.25
104	108.5	154.5	46	2116
105	108.5	27.5	81	6561
106	108.5	196.5	88	7744
107	108.5	170.5	62	3844
108	108.5	68.5	40	1600
109	108.5	89.5	9	81
110	108.5	17	91	8281
111	108.5	148	39.5	1560.25
112	108.5	99.5	9	81
113	108.5	182	73.5	5402.25
114	108.5	184	75.5	5402.25
115	122	131	9	81
116	122	191	69	4761
117	122	179	57	3249
118	122	99.5	22.5	506.25
119	122	22	100	10000
120	122	116	6	36
121	122	154.5	32.5	1056.25
122	122	22	100	10000
123	122	27.5	94.5	8930.25

TABLE 2

(Continued)

Pupil	A	B	D	D^2
124	122	154	32	1024
125	122	99.5	22.5	506.25
126	122	99.5	22.5	506.25
127	122	13	109	11881
128	122	154.5	32.5	1056.25
129	122	79.5	42.5	1806.25
130	133	99.5	33.5	1122.25
131	133	191	58	3364
132	133	141	8	64
133	133	163	30	900
134	133	173	40	1600
135	133	141	8	64
136	133	163	30	900
137	140	163	23	529
138	140	163	23	529
139	140	89.5	50.5	2550.25
140	140	184	44	1936
141	140	108	32	1024
142	140	186.5	46.5	2162.25
143	140	131	9	81
144	147.5	194.5	47	2209
145	147.5	79.5	68	4624
146	147.5	179	32.5	1056.25
147	147.5	89.5	58	3364
148	147.5	188.5	41	1681
149	147.5	68.5	79	6241
150	147.5	40	107.5	11556.25
151	147.5	116	29.5	870.25
152	155	196.5	41.5	1722.25
153	155	191	36	1296
154	155	116	39	1521
155	155	163	8	64
156	155	99.5	55.5	3080.25
157	155	154.5	.5	.25
158	155	148	7.	49
159	161.5	62.5	99	9801
160	161.5	99.5	62	3844
161	161.5	116	45.5	2070.25
162	161.5	62.5	99	9801
163	161.5	116	45.5	2070.25
164	161.5	186.5	25	625

TABLE 2

(Continued)

Pupil no.	A	B	D	D^2
165	169.5	148	21.5	462.25
166	169.5	154.5	15	225
167	169.5	108	61	3721
168	169.5	79.5	90	8100
169	169.5	175	5.5	30.25
170	169.5	79.5	90	8100
171	169.5	179	10.5	110.25
172	169.5	163	6.5	42.25
173	169.5	175	5.5	30.25
174	169.5	182	12.5	156.25
175	176.5	123	53.5	2862.25
176	176.5	141	35.5	1260.25
177	176.5	62.5	114	12996
178	176.5	68.5	108	11664
179	181.5	131	50.5	2550.25
180	181.5	184	2.5	6.25
181	181.5	68.5	113	12769
182	181.5	123	58.5	3422.25
183	181.5	198	17.5	306.25
184	181.5	175	6.5	42.25
185	186.5	170.5	16	256
186	186.5	99.5	87	7569
187	186.5	163	23.5	552.25
188	186.5	116	70.5	4970.25
189	189.5	170.5	19	361
190	189.5	188.5	1	1
191	193	141	52	2704
192	193	131	62	3844
193	193	131	62	3844
194	193	89.5	103.5	10709.25
195	193	194.5	1.5	2.25
196	198	123	75	5625
197	198	199	1	1
198	198	170.5	27.5	756.25
199	198	200	2	4
200	198	141	57	3249
				485474.25

$$\rho = 1 - \frac{6 \sum D^2}{n(n^2 - 1)}$$

$$= .64$$

$$r = .65$$

TABLE 3

RANKING CORRELATION BETWEEN TEACHERS' RATING AND ESSAY-TYPE TEST

Pupil No.	A	B	D	D^2
1	3	59	56	3481
2	3	6	3	9
3	3	170.5	167.5	28129.25
4	3	148	145	21025
5	3	41	38	1444
6	12	61.5	49.5	2450.25
7	12	58	36	1296
8	12	18.5	6.5	42.25
9	12	117	105	11236
10	12	108.5	96.5	9312.25
11	12	148	136	18496
12	12	18.5	6.5	42.25
13	12	1	1	
14	12	18.5	6.5	42.25
15	12	24.5	12.5	156.25
16	12	4	8	64
17	12	67.5	55.5	3080.25
18	12	92.5	80.5	6480.25
19	22	44	22	484
20	22	75.5	53.5	2862.25
21	22	31	9	81
22	22	11	11	121
23	22	31	9	81
24	22	55	33	1089
25	22	24.5	2.5	6.25
26	31	113	82	6724
27	31	14.5	16.5	272.25
28	31	178	147	21609
29	31	187.5	154.5	23956.25
30	31	174.5	143.5	20539.25
31	31	126.5	95.5	9120.25
32	31	18.5	12.5	156.25
33	31	55	23	529
34	31	47.5	16.5	272.25
35	31	18.5	12.5	156.25
36	31	43	12	144
37	41.5	64	22.5	506.25
38	41.5	24.5	17	289
39	41.5	89.5	48	2304
40	41.5	4	37.5	1406.25
41	41.5	11	30.5	930.25

TABLE 3

(Continued)

Pupil No.	A	B	D	D^2
42	41.5	75.5	34	1156
43	41.5	75.5	34	1156
44	41.5	38	2.5	6.25
45	41.5	31	10.5	110.25
46	41.5	51.5	10	100
47	51	126.5	75.5	5700.25
48	51	44	7	49
49	51	8	47	2209
50	51	1.5	49.5	2450.25
51	51	61.5	10.5	110.25
52	51	61.5	10.5	110.25
53	51	18.5	32.5	1056.25
54	51	51.5	.5	.25
55	62	96	34	1156
56	62	187.5	125.5	15775.25
57	62	121	59	3481
58	62	103.5	41.5	1722.25
59	62	61.5	.5	.25
60	62	89.5	27.5	756.25
61	62	89.5	27.5	756.25
62	62	155	93	8649
63	62	153	91	8281
64	62	24.5	37.5	1406.25
65	62	24.5	37.5	1406.25
66	62	14.5	47.5	2256.25
67	62	36	26	676
68	73	36	37	1369
69	73	153	80	6400
70	73	129	56	3136
71	73	96	23	529
72	73	129	56	3136
73	73	31	42	1764
74	73	103.5	30.5	930.25
75	73	67.5	5.5	30.25
76	79.5	196.5	117	13689
77	79.5	41	37.5	1406.25
78	79.5	47.5	32	1024
79	79.5	113	34.5	1190.25
80	87	141	54	2916
81	87	178	91	8281
82	87	187.5	100.5	10050.25
83	87	182	95	9025

TABLE 3

(Continued)

Pupil No.	A	B	D	D^2
84	87	132.5	45.5	2070.25
85	87	157	70	4900
86	87	157	70	4900
87	87	51.5	35.5	1260.25
88	87	36	51	2601
89	87	8	79	6241
90	87	31	56	3136
91	97.5	174.5	77	5929
92	97.5	162.5	65	4225
93	97.5	178	80.5	6480.25
94	97.5	167	69.5	4830.25
95	97.5	75.5	22	484
96	97.5	31	66.5	4422.25
97	97.5	81	16.5	272.25
98	97.5	4	93.5	8742.25
99	97.5	121	23.5	552.25
100	97.5	47.5	50	2500
101	108.5	138.5	30	900
102	108.5	187.5	79	6241
103	108.5	129	20.5	420.25
104	108.5	113	4.5	20.25
105	108.5	148	39.5	1560.25
106	108.5	71	37.5	1406.25
107	108.5	47.5	61	3721
108	108.5	121	12.5	156.25
109	108.5	55	53.5	2862.25
110	108.5	24.5	84	7056
111	108.5	162.5	54	2916
112	108.5	67.5	41	1681
113	122	124.5	2.5	6.25
114	122	167	45	2025
115	122	132.5	10.5	110.25
116	122	153	31	961
117	122	83	39	1521
118	122	143.5	21.5	462.25
119	122	31	91	8281
120	122	67.5	54.5	2970.25
121	122	113	9	81
122	122	67.5	54.5	2970.25
123	122	103.5	18.5	342.25
124	122	89.5	32.5	1056.25
125	122	81	41	1681

TABLE 3

(Continued)

Pupil No.	A	B	D	D^2
126	122	96	26	676
127	122	96	26	676
128	133	170.5	36.5	1332.25
129	133	103.5	29.5	870.25
130	133	47.5	85.5	7310.25
131	133	162.5	39.5	1560.25
132	133	113	120	14400
133	133	103.5	29.5	870.25
134	133	198.5	65.5	4290.25
135	140	1.5	138.5	19204.25
136	140	8	132	17424
137	140	85.5	64.5	4160.25
138	140	103.5	36.5	1332.25
139	140	81	59	3481
140	140	75.5	74.5	5550.25
141	140	113	27	729
142	147.5	198.5	51	2601
143	147.5	167	19.5	380.25
144	147.5	126.5	21	441
145	147.5	192	54.5	2970.25
146	147.5	141	6.5	42.25
147	147.5	75.5	72	5184
148	147.5	121	26.5	702.25
149	147.5	148	.5	.25
150	73	55	18	324
151	155	200	45	2025
152	155	198.5	43.5	1892.25
153	155	190.5	34.5	1190.25
154	155	148	7	49
155	155	148	7	49
156	155	85.5	29.5	870.25
157	155	75.5	79.5	6320.25
158	161.5	113	47.5	2256.25
159	161.5	132.5	29	841
160	161.5	190.5	29	841
161	161.5	157	4.5	20.25
162	161.5	67.5	94	8836
163	161.5	75.5	86	7396
164	169.5	174.5	5	.25
165	169.5	108.5	61	3721
166	169.5	162.5	7	49
167	169.5	136	33.5	1122.25

TABLE 3

(Continued)

Pupil No.	A	B	D	D^2
168	169.5	174.5	5	25
169	169.5	132.5	37	1369
170	169.5	159	10.5	110.25
171	169.5	182	22.5	506.25
172	169.5	180	11.5	132.25
173	169.5	129	40.5	1640.25
174	181.5	193.5	12	144
175	181.5	124.5	57	3249
176	176.5	118	58.5	3422.25
177	176.5	136	40.5	1640.25
178	181.5	170.5	11	121
179	181.5	178	3.5	12.25
180	181.5	170.5	11	121
181	181.5	85.5	96	9216
182	181.5	55	126.5	16196.25
183	181.5	162	19.5	380.25
184	186.5	196.5	10	100
185	186.5	138.5	48	2304
186	186.5	195	8.5	72.25
187	186.5	103.5	83	6889
188	189.5	182	7.5	56.25
189	189.5	121	67.5	4556.25
190	193	184.5	8.5	72.25
191	193	92.5	100.5	10050.25
192	193	193.5	.5	.25
193	193	103.5	89.5	8010.25
194	193	162.5	30.5	930.25
195	198	148	50	2500
196	198	184.5	13.5	182.25
197	198	141	57	3249
198	198	85.5	112.5	12644.25
199	198	92.5	105.5	11125.25
200	198	184	14	196

TOTAL 687006.5

$$\rho = 1 - \frac{6\sum D^2}{n(n^2 - 1)}$$

$$= .48$$

$$r = .497.$$

TABLE 4. SHOWING CORRELATION BETWEEN OBJECTIVE AND ESSAY-TYPE TEST

ESSAY - TYPE TEST. DEPARTMENTAL EXAMINATION, 1932, GRADE VIII.

OBJECTIVE-TYPE TEST	5	10	15	20	25	30	35	40	45	50	Fy	Y	FyY	FyY^2
50														
45						2 +6	3 +9	6 +18	3 +9	3 +9	17	+3	51	153
40		1 +2	2 +4	1 +2	3 +6	8 +16	11 +22	7 +14	3 +6		36	+2	72	144
35			2 +2	7 +7	21 +21	24 +24	16 +16	4 +4	7 +7		81	+1	81	81
30	2 0	2 0	7 0	9 0	18 0	31 0	12 0	9 0	3 0		93	0	0	0
25	3 -3	8 -8	5 -5	10 -10	20 -20	17 -17	7 -7	6 -6			76	-1	-76	76
20	3 -6	8 -16	11 -22	5 -10	10 -20	17 -34	5 -10				59	-2	-118	236
15	2 -6	6 -18	2 -6	4 -12	1 -3	2 -6		2 -6			19	-3	-57	171
10	3 -12						1 -4				4	-4	-16	64
Fx	13	25	29	36	75	102	58	31	16		385		-63	925
X	-5	-4	-3	-2	-1	0	1	2	3					
FxX	-65	-100	-87	-72	-75	0	58	62	48		-231	=FxX		
FxX^2	325	400	261	144	75	0	58	124	144		1531	$=FxX^2$		
Σ FxyY	-27	-40	-27	-23	-10	-8	35	15	22		-63	=FxyY		
$X\Sigma FxyY$	135	160	81	46	10	0	35	30	66		563	=FxyYX		

$r = \frac{a}{\sqrt{bc}}$

a = -525.2

b = -1392.4

c = 914.7

r = .46.

TABLE 5

SHOWING CORRELATION BETWEEN OBJECTIVE AND ESSAY-TYPE TESTS GIVEN IN BRANDON, 1933

ESSAY-TYPE TEST \ SHORT ANSWER TEST	1 4	14	24	34	44	54	64	75	Fy	Y	FyY	FyY^2
100												
89					1	2	1	1	5	+4	20	80
79					+4	+8	+4	+4				
69						8	10		18	+3	54	162
						+24	+30					
59				1	6	7	13	1	31	+2	62	124
				+2	+18	+14	+26	+2				
49				2	2	8	11		25	+1	25	25
				+2	+2	+8	+11					
39		1			4	12	8	2	30	0	0	0
		0			0	0	0	0				
29				2	4	12	6		30	-1	-30	30
				-2	-4	-12	-6					
19				3	7	14	5		29	-2	-58	116
				-6	-14	-28	-10					
9			1	1	11	14	2		29	-3	-87	261
			-3	-3	-33	-42	-6					
1			1	4	3	3			11	-4	-44	176
			-4	-16	-12	-12						
Fx	1	2	13	47	80	59		4	206		-60	972
X	-4	-3	-2	-1	0	+1		+2				
FxY	-4	-6	-26	-47	0	59		8	-16	208		
FxY2	16	18	52	47	0	59		16	208			
$\Sigma FxyY$	0	-7	-23	-42	-40	+46		+6	-60			
$\Sigma FxyY$	0	21	46	42	0	46		12	167			

$$r = \frac{a}{\sqrt{bc}}$$

a = 162.34
b = 206.76
c = 954.52

r = .36

TABLE NO: 6. SHOWING CORRELATION BETWEEN OTIS TEST AND ESSAY-TYPE TEST

ESSAY — TYPE TEST \ OTIS	20	25	30	35	40	45	50	55	60	65	70	75	Fy	Y	FyY	FyY²
90																
85																
80					1 +7				1 +7	2 +14			4	+7	28	196
75									1 +6	1 +6	1 +6	1 +6	4	+6	24	144
70								1 +5		2 +10	5 +25	1 +5	9	+5	45	225
65					1 +4		1 -4	3 +12	2 +8	8 +32	4 +16		19	+4	76	304
60								1 +3	3 +9	3 +9	2 +6	1 +3	10	+3	30	90
55							2 +4				4 +8	1 +2	7	+2	14	28
50			1 -1						1 +1	2 +2	3 +3		7	+1	7	7
45	0				1 0		2 0	4 0		2 0	2 0	1 0	12	0	0	0
40			1 -1						4 -4	1 -1	1 -1	1 -1	8	-1	-8	8
35					1 -2	1 -2	3 -6	4 -8	3 -6	2 -4	1 -2		15	-2	-30	60
30						1 -3		2 -6	1 -3	3 -9	3 -9		10	-3	-30	90
25								4 -16	1 -4	5 -20	1 -4		11	-4	-44	176
20								4 -20	3 -15	3 -15			10	-5	-50	250
15				2 -12	3 -18	3 -18		1 -6	5 -30	2 -12	2 -12		18	-6	-108	648
10					2 -14	1 -7		3 -21	2 -14			1 -7	9	-7	-63	441
5								2 -16	2 -16				4	-8	-72	256
Fx		2	0	5	8	12		29	29	76	29	7	157		-141	2922
X		-6	-5	-4	-3	-2		-1	0	+1	+2	+3				
FxX		-12	0	-20	-24	-24		-29	0	76	58	21	6			
FxX²		72	0	80	72	48		29	0	76	116	63	51	6		
FxyY		0	0	-10	-30	-23		-7	-61	-12	36	8	-141			
X≤FxyY		0	0	40	90	46		73	0	12	72	24	757			

$$r = \frac{a}{\sqrt{bc}}$$

a= 762.78

b= 813.78

c= 2796.87

r= .80

TABLE 7. SHOWING CORRELATION OF OBJECTIVE TEST WITH OTIS TEST

Objective \ Otis	20	25	30	35	40	45	50	55	60	65	70	Fy	y	Fyy	Fyy²
75															
70															
65											2 +6	2	+3	6	18
60									4 +8	5 +10	1 +2	10	+2	20	40
55							4 +4	2 +2	8 +8	15 +15	4 +4	33	+1	33	33
50					2 0	3 0	5 0	7 0	10 0	12 0	2 0	41	0	0	0
45			3 -3	1 -1	1 -1	5 -5	12 -12	10 -10	1 -1			33	-1	-33	33
40	1 -2		1 -2	4 -8	5 -10	10 -20	6 -12					27	-2	-54	108
35			1 -3	1 -3	2 -6	4 -12						8	-3	-24	72
30		1 -4		1 -4	1 -4							3	-4	-12	48
Fx	1	1	5	9	12	28	27	32	33	9		157		-64	352
x	-5	-4	-3	-2	-1	0	+1	+2	+3	+4					
Fxx	-5	-4	-15	-18	-12	0	27	64	99	36		172			
Fxx²	25	16	45	3	12	0	27	128	297	144		730			
Σfxyy	-2	-4	-8	-16	-21	-33	-22	6	24	12		-64			
xΣfxyy	10	16	24	32	21	0	22	12	72	48		257			

$$r = \frac{a}{\sqrt{bc}}$$

a = 327.1

b = 541.6

c = 325.9

r = .78

TABLE 7. SHOWING CORRELATION OF OBJECTIVE TEST WITH OTIS TEST

TABLE 9.

SHOWING CORRELATION BETWEEN PLANNED-PARAGRAPH TEST AND OTIS TEST

OTIS TEST \ PLANNED-PARAGRAPH TEST	15–19	20–24	25–29	30–34	35–39	40–44	45–49	50–54	55–59	60–64	65–69	70–74	75–79	80–84	Fy	Y	FyY	FyY^2
95–100										+5 [1]					1	+5	5	25
90–94							+4 [1]			+8 [2]	+4 [1]	+4 [1]			5	+4	20	80
85–89												+9 [3]			3	+3	9	27
80–84								+6 [3]		+2 [1]	+8 [4]	+2 [1]			9	+2	18	36
75–79							+3 [3]	+2 [2]	+4 [4]	+1 [1]		+1 [1]			11	+1	11	11
70–74				0 [1]	0	0 [1]	0 [1]	0 [1]	0	0 [1]	0 [2]	0 [1]	0 [1]		9	0	0	0
65–69				-1 [1]	-3 [3]	-1 [1]	-2 [2]	-2 [2]	-2 [2]	-1 [1]		-2 [2]			14	-1	-14	14
60–64	-2 [1]			-2 [1]		-2 [1]	-10 [5]	-4 [2]	-2 [1]		-2 [1]				12	-2	-24	48
55–59	-3 [1]			-3 [1]	-3 [1]										3	-3	-9	27
50–54		-4 [1]													1	-4	-4	16
45–49					-5 [1]										1	-5	-5	25
40–44					-6 [1]										1	-6	-6	36
Fx	2	1	0	4	6	4	12	10	7	7	7	9	1		70		1	345
X	-6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6					
FxX	-12	-5	0	-12	-12	-4	0	10	14	21	28	45	6		79			
FxX^2	36	25	0	36	24	4	0	10	28	63	112	225	36		599			
$\Sigma FxyY$	-5	-4	0	-6	-17	-3	-5	2	0	15	10	14	0		1			
$X\Sigma FxyY$	30	20	0	18	34	3	0	2	0	45	40	70	0		262			

$$r = \frac{a}{\sqrt{bc}}$$

$a = 260.87$

$b = 509.84$

$c = 344.99$

$r = .62$

TABLE 10.

SHOWING CORRELATION BETWEEN PLANNED-PARAGRAPH TEST AND SHORT-ANSWER TEST

SHORT ANSWER TEST	PLANNED PARAGRAPH TEST: 15–19	20–24	25–29	30–34	35–39	40–44	45–49	50–54	55–59	60–64	65–69	70–74	75–79	Fy	Y	FyY	FyY^2
90–95												$+4^{1}$		1	+4	4	16
85–89						$+3^{1}$		$+6^{2}$	$+3^{1}$	$+9^{3}$		$+3^{1}$		8	+3	24	72
80–84					$+4^{2}$		$+6^{3}$	$+2^{1}$	$+6^{3}$	$+4^{2}$	$+2^{1}$	$+6^{3}$	$+2^{1}$	16	+2	32	64
75–79				$+1^{1}$	$+1^{1}$		$+2^{2}$	$+3^{3}$	$+2^{2}$	$+3^{3}$	$+1^{1}$	$+3^{3}$	$+1^{1}$	17	+1	17	17
70–74				0^{3}	0^{1}	0^{1}		0^{2}	0		0^{1}			8	0	0	0
65–69		-1^{1}			-2^{2}		-2^{2}			-1^{1}				6	-1	-6	6
60–64				-2^{1}		-2^{1}		-4^{2}		-4^{2}	-2^{1}			7	-2	-14	28
55–59	-3^{1}						-3^{1}							2	-3	-6	18
50–54								-4^{1}						1	-4	-4	16
45–49	-5^{1}				-10^{2}									3	-5	-15	75
40–44	-6^{1}													1	-6	-6	36
Fx	3	1	0	5	8	3	8	11	6	11	4	8	2	70		26	348
X	-7	-6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5				
FxX	-21	-6	0	-20	-24	-6	-8	0	6	22	12	32	10	-3			
FxX^2	147	36	0	80	72	12	8	0	6	44	36	128	50	619			
$\Sigma FxyY$	-14	-1	0	-1	-7	+1	+3	+3	+11	+11	+1	16	3	26			
$X\Sigma FxyY$	98	6	0	4	21	-2	-3	0	11	22	3	64	15	239			

$$r = \frac{a}{\sqrt{bc}}$$

a= 240.11

b= 618.87

c= 338.34

r= .60

TABLE 10. SHOWING CORRELATION BETWEEN PLANNED-PARAGRAPH TEST AND SHORT-ANSWER TEST

TABLE 11.

TABLE OF FINDINGS OBTAINED FROM STATISTICAL STUDY OF SCORES.

	Objective Test	Essay-Type Test	Planned Paragraph Test
Average	50.92	42.29	
Reliability Coefficient	.92	.68	
S. D.	7.20	21.10	13.1
P. E. est.	1.94	10.38	
Correlation with Otis	.81	.32	.62
Correlation with Teachers' Rating	.67	.49	