

THE UNIVERSITY OF MANITOBA

A COMPARATIVE STUDY OF COLLEGIATE
AND DEPARTMENT OF EDUCATION
MARKS IN MANITOBA

BEING A THESIS SUBMITTED TO THE COMMITTEE
ON POST-GRADUATE STUDIES IN PARTIAL
FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF
EDUCATION

BY

WALLACE HENRY McINTYRE

WINNIPEG, MANITOBA

APRIL, 1952



ACKNOWLEDGEMENTS

The writer wishes to express his sincere appreciation to the Principals of Gordon Bell, Kelvin, West Kildonan, Dauphin, Morden, Portage la Prairie and Glenlawn Collegiates for their co-operation in obtaining data used in this investigation; to Dr. D.S.Woods, former Dean of the Faculty of Education, for his guidance and help throughout the period in which the study was conducted.



TABLE OF CONTENTS

	Page
LIST OF TABLES	iv v
LIST OF FIGURES	vi
CHAPTER	
I. INTRODUCTION	I
Measurement by Marks	
Secondary School Examination	
System in Manitoba	
Purpose of this Study	
Sources of Data	
School Marks	
Department of Education Marks	
II. HISTORY AND PLACE OF MARKS IN EVALUATING	
EDUCATIONAL PROGRESS	17
Brief Review of the Early History	
of Examinations	
Period of First Scientific Measurement	
in Education	
Progress Since 1918	
III. A STUDY OF STUDENT MARKS IN SIX	
COLLEGIATES FOR 1949	35
Gordon Bell	
West Kildonan	
Morden	
Kelvin	
Dauphin	
Portage la Prairie	
IV. COMPARATIVE STUDY FOR THE YEAR 1950 . . .	49
V. THIRD YEAR OF INVESTIGATION, 1951 . . .	65
VI. INTERPRETATION OF CUMULATIVE DATA . . .	81
VII. CONCLUSION	93

TABLE OF CONTENTS (CONT'D)

APPENDIX

	Page
A. GORDON BELL AND DEPARTMENT MARKS FOR 1949	96
KELVIN " " " " "	98
WEST KILDONAN " " " " "	99
DAUPHIN " " " " "	100
MORDEN " " " " "	100
PORTAGE LA PRAIRIE "	101
B. GLENLAWN(ST.VITAL) AND DEPARTMENT MARKS	
FOR 1949, 1950 and 1951	102
BIBLIOGRAPHY	103

LIST OF TABLES

Table	Page
1. Number of Students	12
2. Summary of Gordon Bell Examination Results compared with Department of Education Marks for 1949	36
3. Summary of Kelvin Examination Results compared with Department of Education Marks for 1949	37
4. Summary of West Kildonan Examination Results compared with Department of Education Marks for 1949	38
5. Summary of Dauphin Examination Results compared with Department of Education Marks for 1949	39
6. Summary of Morden Examination Results compared with Department of Education Marks for 1949	40
7. Summary of Portage la Prairie Examination Results compared with Department of Education Marks for 1949	41
8. Range in Means of School Marks compared with the Range in Department of Education Marks, 1949	47
9. Summary of Gordon Bell Examination Results compared with Department of Education Marks for 1950	49
10. Pupil Variation in Wide Range Subjects in Gordon Bell, 1950	50
11. Summary of Kelvin Examination Results compared with Department of Education Marks for 1950	51
12. Pupil Variation in Wide Range Subjects in Kelvin, 1950	52
13. Summary of West Kildonan Examination Results compared with Department of Education for 1950	53
14. Pupil Variation in Wide Range Subjects in West Kildonan, 1950	54
15. Summary of Dauphin Examination Results compared with Department of Education Marks for 1950	55
16. Pupil Variation in Wide Range Subjects in Dauphin, 1950	55
17. Summary of Morden Examination Results compared with Department of Education Marks for 1950	56

LIST OF TABLES
(continued)

Table	Page
18. Pupil Variation in Wide Range Subjects in Morden, 1950	57
19. Summary of Portage la Prairie Examination Results compared with Department of Education Marks for 1950	57
20. Pupil Variation in Wide Range Subjects in Portage la Prairie, 1950	58
21. Subjects in Collegiates showing Significant Over-Eval.	59
22. Subjects in Collegiates showing Significant Under-Evaluation	59
23. Range in Means of School Marks compared with the range of Department of Education Marks, 1950	63
24. Summary of Gordon Bell Examination Results compared with Department of Education Marks for 1951	65
25. Summary of Kelvin Examination Results compared with Department of Education Marks for 1951	66
26. Summary of West Kildonan Examination Results compared with Department of Education Marks for 1951	67
27. Summary of Dauphin Examination Results compared with Department of Education Marks for 1951	68
28. Summary of Morden Examination Results compared with Department of Education Marks for 1951	69
29. Summary of Portage la Prairie Examination Results compared with Department of Education Marks for 1951	70
30. Number of Subjects Outside Five-Point Range	71
31. Range in Means in School Marks compared with the Range of Means in Department of Education Marks, 1951	76
32. Comparative Analysis of Ranges in Means for Collegiate Marks and Department of Education Marks	89
33. Summary of Equal Scores in School Marks and Department of Education Marks for the Three Years	91

LIST OF FIGURES

Figure	Page
1. Variation in Subject Averages for 1949	43
2. Variation in Subject Averages for 1949	44
3. Composite Graph on Variation for 1949	45
4. Composite Graph on Variation for 1950	61
5. Composite Graph on Variation for 1951	74
6. Percentage of Subjects for Three Years	86
7. Percentage of Subjects for Three Years	87

CHAPTER I
INTRODUCTION

Marks play such a significant role in our educational system that they are a never-ending field for investigation. From the time a child enters school, marks are with him until he graduates. They not only indicate present progress in school work but continue to affect a child, directly or indirectly, all his life, and have, consequently, a powerful social influence. Some system of measurement is essential in order to present, in a manner that indicates as exactly as possible, pupil progress. Learning involves psychological processes and it is impossible for any system to measure the working of the human mind with the accuracy that can be secured in measuring human efficiency in limited operations such as cutting down a tree, or measuring a quart of milk. As the secondary school becomes more complicated and demanding, the problem of establishing a highly valid system of measurement acquires ever increasing importance. Such a system must not only be valid among educational bodies, but must also be acceptable in the economic world into which every child is eventually absorbed. The present marking system, is the answer presented by our educators to the need for measurement, and, accordingly, is subjected to constant study with a view to rendering it more adequate.

Controversy has been endless in the study of this educational instrument, the effort always being towards the establishment of potential accuracy and comparative validity. The present investigation employs what the writer believes to be a worthwhile approach to the evaluation of the system in use at present in the Province of Manitoba.

Measurement by Marks

Marks are used extensively in all secondary schools throughout the academic year, and contribute to the establishment of a final mark at the end of the school year. This final mark carries great weight in that it is accepted as a measure of progress throughout the year. Teachers' marks, therefore, occupy a vital position in securing uniformity in estimating pupil progress across the Province as a whole.

In order to focus attention upon the importance of marks, it is important to have a general understanding of their origin and development. Two phases of this subject are evident, namely, (1) origin and purpose, (2) the effect of the changing conception of education in reference to the marking system. Let us consider first the original purpose of examinations. Briefly, examinations formerly endeavored to test the mastery of certain narrow fields of subject-matter. Naturally, the teaching was as simple as the system of testing, and consisted of definite assignments by the teacher, study of them by the pupil, recitation or tests of assignments in class, and, lastly, the

evaluation of learning by means of marks obtained in a written examination. As will be seen later, recent years have brought noteworthy changes in this educational program of the secondary school. Secondly, as the concept of education has changed, the method of evaluation has changed accordingly. The mastery of subject-matter as the chief purpose of education has given way to a much broader goal which emphasizes the achievement of well-rounded pupil personality. The two developments are closely interlocked, and, as the technique of instruction has changed, so the purpose of examinations has broadened. Marking, as an accurate representation of student intellectual growth, has inevitably been profoundly affected.

The following two quotations will reveal something of the old and the new conceptions of marks as the measure of intellectual growth:

"Traditionally, the sole basis for marks has been subject matter achievement as measured by tests and examinations. This was only natural, since until recently the mastery of subject matter was the chief immediate goal of the secondary school program. The pupil was expected to accumulate a designated mass of facts and information. The mark was used to indicate the degree to which he succeeded. The inadequacy of subject matter as the sole basis for marks has long been recognized. Obviously, such a policy is not at all in harmony with the philosophy that presently underlies the program of educational units." ¹

¹ W.T.Gruhn-H.R.Douglas. The Modern Junior High School. New York: The Ronald Press Co. 1947, p.385.

"Since the teacher's mark is now and almost certainly will continue to be the most frequently recorded measure of pupil accomplishment, it is very important that the class room teacher have a definite notion of the functions of such marks. Teacher's marks function in a four-fold way: (1) they provide the basis for the school's record of the child's educational history, (2) they furnish the teacher with a record of the pupil's achievement and progress, (3) they reveal to the pupil the school's evaluation of his effort and accomplishments, and (4) they furnish to the parent reasonably accurate information concerning the pupil's achievement. For school administrators marks afford the common basis for determination of promotions, scholastic honors, and school classification. For the teacher, marks provide a working basis for group distinctions in assignments, work requirements, extra-curricular activities, etc. For the pupil, marks should give accurate information concerning the amount and quality of work done. For the parent, the marking system should supply accurate information on pupil achievement which should indicate relative success or failure in unmistakable terms. Obviously, the realization of these four-fold functions of the marking system places it under a very severe burden. The real severity of this burden is better appreciated when we recall the implications of the experimental evidence on the reliability of teachers' marks, and then in the face of these disturbing facts, realize the seriousness with which these marks are taken by the pupil, by the parents, and even by the school itself." 2

In the Canadian Provinces marks are the measure for admittance to higher courses of study. Hence, the centralized, provincial examination for all pupils not accredited by the particular school is necessitated.

2

H.A. Greene-A.N. Jorgensen. The Use and Interpretation of High School Tests. New York: Longmans, Green and Co. 1940, p. 566.

Secondary School Examination System in Manitoba

Bearing in mind, then, how important it is that the final score should be accepted as truly representative of the year's progress, satisfactory evidence must be available that, according to some general standard of measurement, it is as close to being accurate as possible. The standard in Manitoba is the one set by the Department of Education in a series of examinations.

The final mark determined at the end of the school year should compare closely with this generally accepted standard. This standard, as noted, is established by a series of examinations set and marked under the supervision of the Department of Education in June and July, and written by certain groups of students to whom we shall later refer. Accordingly, class marks, as determined by teachers in individual schools, can be checked against those obtained in Departmental examinations and their worth estimated.

The examination system directed by the Department of Education to-day is the result of many years of sustained effort and constant development. As in the past, the Department is the administrative centre of examinations, but authority is not now as narrowly centralized as formerly. Fifty years ago, the examination system was dominated by the Department of Education for all grades from VIII to XII inclusive. These examinations were held throughout the Province. Committees,

picked by the Advisory Board, set these Departmental examinations. All written papers were marked in Winnipeg by committees appointed jointly by the Department of Education and the University. The marks given were accepted by both the Department and the University. There was, however, one important difference in acceptance of these marks. In some subjects, a different pass standard was set by the Department and by the University. This meant that each educational authority was a law unto itself and each prescribed its own standards. If a student failed in two or more subjects, he had to repeat his year. The pass standards generally were high, and Manitoba was considered to have a satisfactory system of examining.

Shortly after the First World War, Grade VIII examinations directed by the Department, were abolished. About the same time another innovation was made. The number of examinations was reduced in Grade IX to four subjects—those subjects which were not carried on past Grade IX. In Grade X, examinations were reduced to those subjects that were completed in Grade X and not carried on into Grade XI. The Grade XI examinations completed the roster. Gradually the number of places approved for the writing of Department of Education examinations was increased. The pass standard was fixed at 50 but, as before, if there were two failures, all the subjects had to be repeated. Later final credit was allowed for every 50% subject and only those below 50% had to be repeated. The University still compiled and interpreted its own set of marks and the Department retained

its own.

This dual operation continued until the middle thirties when the Manitoba Examination Board, consisting of University and Department representatives, was established. This is now a statutory body controlling the setting of examinations. When a Departmental paper is set, University and Department representatives work as a committee to produce a suitable examining paper that is acceptable to both authorities. The Department representatives are selected from teachers actually engaged in teaching the subject being examined, so that they are thoroughly cognizant of both the student and the teacher problems. To-day there is one set of marks, which is retained in the Department, as the University now accepts this joint marking. Changes may occur in the curriculum from time to time but the examination papers are adjusted accordingly.

Purpose of this Study

We have indicated in the foregoing sections the foundation for our study, namely, establishment of the fact that the Department of Education is the central authority for examinations in Manitoba. Schools throughout the Province with the same official standing have students writing the same examinations all of which are marked by the same marking committees. In addition all schools record teachers' and general school examination marks. With these lists of marks as our data, the investigation begins. Its purpose is to determine whether a comparatively close, or a very wide,

discrepancy exists between school marks and those of the Department of Education. In other words, do these two sets of marks reveal marked over-estimation or under-estimation when we compare the school marks with the Departmental marks? An illustration may clarify our position. If a school gives a student 50 in a subject and he makes 60 in the Departmental examination, the school mark is lower by 10 and the net result is scored as a -10. On the other hand, if the student makes 35 on the Department examination, he would be scored as +15. Detailed records and tables referring to this point will be found in succeeding chapters.

In the course of this study, which is conducted through schools of the same rating for a period of three consecutive years, 1949, 1950, and 1951, the writer aims to discover any local trends that affect marks. Should a school reveal significant over-evaluation in a subject, a tendency to over-estimation exists. This phase of the study will be expanded as the investigation develops.

Finally, as a secondary deduction developing from the major aim of the investigation, the writer endeavours to explore the probability that certain subjects by their inherent nature possess advantages over others from the point of view of exactness in marking. Should Composition produce closer scores than French? This, and other observations of similar nature, will develop as the study proceeds. The main purpose, however, of this investigation is to determine the existence of tendencies towards under-estimation or over-

estimation of school marks as compared to those awarded on the Provincial examination.

Sources of Data

All the data for this study are obtained from two main sources:

- (1) Record Sheets of Accredited Collegiates.
- (2) Official Records of the Department of Education.

Before discussing these sources, let us examine the meaning of the word "accredited" as applied to schools in Manitoba.

Accredited Schools; The following quotation indicates the type of school involved:

"Each Collegiate Institute, on being granted powers of recommending pupils without Departmental Examinations shall be classified as an 'Accredited Collegiate Institute'. . . . Pupils who are not recommended may write the Departmental Examinations. . . . and must obtain a standing on the year's work of not less than fifty per cent (50%) in any subject and an average of at least sixty-seven (67%) on the full Second Level (Grade XI)." ³

The power to recommend pupils without Department examinations, i.e., the standing as an Accredited Collegiate, is granted to a Collegiate when it can meet the regulations set down by the Department in three respects: (1) teachers' and principals' academic qualifications and experience, (2) proper equipment for laboratory and library work, (3) certification that pupils have completed required courses in all subjects and attended regularly. In the following

study, students from accredited schools only are included, as the standards governing instruction in these schools are those required by the Department.

A large percentage of the Collegiates in the Province, which could have the privilege of granting accredited standings, do not ask for it. There are seventeen Accredited Collegiates and, for our purpose, we have chosen six, so feel the study should be representative. The City of Winnipeg is represented by Gordon Bell and Kelvin, the suburban schools of Manitoba by West Kildonan (Centennial), and the Province at large by Portage la Prairie, a small city, and two country towns, Morden (Maple Leaf), and Dauphin. There was no thought as to any special selection except in so far as an effort was made to choose types representative of different population areas. An additional set of marks from a second suburban school in St. Vital (Glenlawn) is included in the Appendix for further comparative or reference purposes, if required.

In all these schools, as we have indicated, pupils are recommended on the basis of a high mark standing, and those not receiving 67% average have the opportunity of writing the Department examinations in June. The system has been accepted as fair and is seldom questioned. Most of the students in these schools fulfil the 67% average requirement for recommendation. The others, those below the 67% average, are "non-exempt" and, to obtain their Second Level standing, must

write the examinations of the Department. It is the marks of these students with which we are concerned in this investigation. In Literature, Composition, History, Geometry, Algebra, Chemistry, Physics, French and Latin, the two sets of marks, Collegiate and Department, are studied with a view to determining whether there is significant over or under-scoring by teachers.

1. School Marks; Our first source of data was that secured from school records. School marks have been filed on Kardex Forms, in files, or in suitable books, from year to year, so that the writer was able to secure students' marks quickly and accurately.

In order to have the data in convenient form for later analysis, a Tally Sheet was prepared. On it were recorded the school, the year, and, in four columns, (1) the number of the student, (2) school marks, (3) Departmental marks and (4) a column for purpose of further analysis.

2. Department of Education Marks; The results of the June Examinations are sent by the Department of Education to all schools towards the latter part of July. These scores are contained on special Departmental Forms and become the property of the school. These records of marks in schools date back many years. The writer had access to all such records and it was a simple matter to list them on our Tally Sheets. A total of 3552 marks from each source, the Department and the school

records, was made available from the schools chosen. This number represents approximately 40% of the total papers written. The following table gives the number of students whose marks were analyzed in each Collegiate:

TABLE I
NUMBER OF STUDENTS

School	1949	1950	1951	Total
Gordon Bell	77	59	95	231
Kelvin	72	65	52	189
West Kildonan	16	18	19	53
Dauphin	34	16	30	80
Morden	10	16	8	34
Portage la Prairie	18	17	21	56
Total	227	191	225	<u>643</u>

Method of Estimate of Mark Variation

A few years ago, Mr. O.V. Jewitt, Principal of Gordon Bell High School, in a discussion on the problem of variation, expressed the wish that study could be made into the over-estimation and under-estimation of marks in examinations. He felt that many teachers under-valued marks in order to obtain better efforts from students. In addition, he thought that different subjects inherently produce wide variation in marks. In view of this discussion, the writer undertook a study of Grade XI Matriculation results beginning from October, 1949, at the Gordon Bell school. The staff consisted of teachers with accepted academic standards and a wholesome attitude towards the student body. Their marks could well be used to

establish certain yardsticks for measurement and, with this series as a starting point, the investigation was carried to a number of Collegiates throughout the Province.

In Gordon Bell, as in all Collegiates, the teacher of a subject is responsible for the final marks which are based on the year's work. Special sheets for each class are set up in the school office and teachers must record their marks. The class teacher is responsible for obtaining averages after all marks are listed. It is to be remembered that these are the final school marks and, however they may be determined, the investigation concerns itself only with final scores.

It may assist the reader to repeat, at this point, the method of estimating variation in examination results as between school marks and those of the Department. For example, if a student receives 65 in Literature as his final school mark, and then obtains 70 in June, the school has underestimated him by 5 marks, or the variation is -5. It is to be observed that we are using the Department of Education examination mark as the standard against which to estimate the extent of variation in teachers' marks. If we consider the Literature paper of another pupil with a mark of 65 in school and 55 for the Department, a difference of 10 is revealed and this is scored as +10. Variability is the extent to which the marks spread above or below the Departmental standard.

Tally Sheets are filled out as previously indicated.

The following example illustrates their use:

Name of School

Year

Subject

Student	School Mark	Department Mark	Variation Plus or Minus
<u>1.</u>	<u>76</u>	<u>70</u>	<u>6</u>

Totals on these Tally Sheets are next computed and set down to be used later in the Summary Tables. An illustration of Summary Tables, from which comparisons are made, is presented herewith:

Subject	Number of Papers	Total Plus or Minus	Per Pupil Average Plus or Minus	Number of Papers Equal Dept.
Literature	68	-82	-1.16	3

The number of papers varies considerably in different subjects even in one year, so that each subject is represented by a plus or minus average in relation to the Department of Education average. This means that for the year mentioned above Literature has -1.16 average out of a total of 68 papers. At the same time the writer determined and listed on this table, the number of papers equal in both sets of marks. Further use of all Summary Sheets with detailed analysis in respect to them will constitute the investigation of succeeding chapters.

Kelvin was the second Winnipeg City school chosen for the study. The marks for both school and Department were

entered at the same time. Later calculations were done on the Kelvin Tally Sheets and Summary Tables.

West Kildonan (Centennial) was chosen as the representative of suburban schools. The same procedure has been followed in this Collegiate as in the previous two schools.

Principals of Dauphin Collegiate and Morden (Maple Leaf) Collegiate were mailed Tally Sheets. The writer travelled to Portage la Prairie where the school records were made available. For each of the five Collegiates the procedure has been the same as that described for Gordon Bell.

The same method of analysis of student marks was conducted for all Collegiates for 1950 and 1951. As the study was begun before the Department of Education set a General Science examination, the writer has purposely omitted that subject throughout. Again, certain optional subjects are omitted, as it is believed necessary to have a subject recorded for three consecutive years.

A point that may arise here is the validity of school examinations and those of the Department. It is proper to state that all examinations are measurements of a sort. Department of Education examinations, however, are the only single measure used over the entire Province, are conducted by committees, and have the added weight of wide application. The study involves several hundred students from Accredited Collegiates representative of accredited schools of the Province so that the comparison should be a fair indication

as to whether the teachers' marks are too high or too low. Lastly, with such a wide distribution among schools of the same academic rating, any non-validity of the examination should be offset in part by the number of pupils examined.

The study is made for each year and then the cumulative results for all three years analyzed.

CHAPTER II
HISTORY AND PLACE OF MARKS IN EVALUATING
EDUCATIONAL PROGRESS

It is to be regretted in a study of this nature that it is not possible to treat such a wide topic exhaustively. Yet the writer feels that even a brief treatment will provide a foundation to understanding the significance of marks in our present educational system.

"In order to determine the fruitfulness or wastefulness of methods of learning and teaching school subjects, it is necessary to evaluate the achievements of pupils as accurately as possible. Furthermore, the successful operation of a school demands an accounting of the work of its pupils.

Marks have been the universal measure of school work. So many problems in the management of a school—credit, failure, promotion, retardation, elimination, graduation, honors, recommendations for positions, indeed the entire scholastic machinery of a school—hinge upon the assignment of marks that it is highly imperative to examine in detail the value, accuracy and reliability of marks." ⁴

What is the importance of marks in a Provincial system of education? To establish this point, it is necessary to view the traditional attitude towards marks as an estimate of progress in learning. Significant change in the evaluating of their need for purposes of promotion must be noted. The present chapter represents an attempt to evaluate changing conceptions of the methods of examination and the

⁴ Daniel Starch. Educational Psychology. New York: The Macmillan Company, 1922, p.426.

use of marks applied therewith. Three periods in establishing the place of marks as a means of measurement, are considered:

- (1) Brief Review of the Early History of Examinations
- (2) Period of First Scientific Measurement in Education
- (3) Progress since 1918

Brief Review of the Early History of Examinations

"Examinations have an ancient origin. Where and when they first appeared is not known. Oral examinations are probably the oldest form, evidences of which in ancient literature are found. In the Old Testament we are told the 'Gileadites, at the passage of the Jordan tested the ability of the Ephraimites to pronounce the word Shibboleth. Any Ephraimite who failed to pronounce the aspirate and answered 'Sibboleth' was slain on the spot. On that fatal day forty-two thousand are reported to have failed in their examinations." ⁵

Examinations, of one type or another, are as old as life itself. Primitive people taught their children through direct imitation, and by tribal ceremonies imparted what is commonly known as "tribal knowledge". Formal testing was unknown and all standards were dictated by the law of survival, the knowledge necessary for existence. If prowess and cunning were not enough, the enemy, whether man or beast, lived and the individual died. The chief, as head of a tribe, was the sole authority and won his place often by superior physical ability and cunning. Custodians of knowledge were next found in the priests or their representatives. At first they were the teachers of the people in matters affecting moral, spiritual or supermatural life but, as time went on, they became the centre

⁵H.A. Greene -A.N. Jorgensen, op.cit., p.40.

of all branches of knowledge. Examinations as we know them were unknown, though tests of heroism and military skill had survived from the earliest times. As society became more complex and more organized, local groups, or state, national and educational branches of life, divided. Gradually a simple form of written examination emerged in each of these branches. Tests were held first of all for the few men seeking public office. The oldest known system of examinations was that used in China for the qualification of men for service in the governmental agencies.

"Written examinations are probably of more recent origin than oral quizzes, but even written examinations date back many centuries. As early as 2200 B.C., China had an elaborate national system of examinations for the purpose of selecting her public officials and these examinations have been known down through the ages for their unusual severity." ⁶

And, again,

"A sociologist attributes the remarkable stability of the Chinese civilization, the oldest culture of any modern nation, to five factors, one of which is her highly organized examination system. It began informally in 225 B.C., and became a definite civil service examination system in 29 B.C. The system, described as being thoroughly democratic, ruthless, invariable, and orthodox, has had profound effects, some good and some bad, not only upon the educational system of China, but also upon her whole civilization." ⁷

In Europe, even prior to the 5th century B.C., the educational system, as distinct from religion, was definitely becoming a force in its own right. In Greece, the great philosophers, Socrates, Plato, and Aristotle, attempted to bring a more liberal conception to the world. Oral examinations

⁶ ibid., p.41.

⁷ C.C. Ross. Measurement in To-day's Schools. New York: Prentice-Hall Company, 1947, p.27.

played an important part in their efforts to test the progress of their pupils.

"Socrates, in a method he has made famous, subjected his pupils to exhaustive and searching questioning. Oral quizzing, Socratic or otherwise, has undoubtedly been a part of classroom procedure from the beginnings of teaching activity- in fact, there have been and still are times when it constitutes practically the whole of the teaching act." 8

Definite standards were being acquired and oral catechism and discussion attempted to check approach to these standards. As arts, sciences, trade and commerce developed, diversified interests required diversified learning and the necessity became evident for the measurements of skill and achievement. Life, of course, was still predominantly military, but definite historical milestones, such as the Olympian games under the Grecian regime, spread of Christianity with its catechetical schools, the basic codification of laws of the early Romans, mark eras when standards of measurement were being invented.

"Christian workers came in contact with the best scholars of the Hellenic learning, and particularly at Alexandria, Athens, and the cities of Asia Minor. The speculative Greek would not be satisfied with the simple, unorganized faith of the early Christians. He wanted to understand it as a system of thought, and asked many questions that were hard to answer. To meet the critical inquiry of learned Greeks, it became desirable that the clergy of the Church, in the East at least, should be equipped with a training similar to that of their critics. As a result there was finally evolved, first at Alexandria, and later at other places in the Empire, training schools for the leaders of the Church.

These schools came to be known as catechetical schools, from their oral method of questioning, and this term was later applied to elementary religious instruction (whence catechism) throughout Western Europe." 9

Each period of History added its share to the progress of Civilization as a whole. Education became a definite requirement for those born to opportunity, and even, in numerous cases, for those of outstanding ability and interest in academic learning. As education became more definitely classified and specialized according to the knowledge of those times, a system of tests gradually played a more and ^{more} important role in measuring attainment. Speaking of early teacher and student guilds during the Middle Ages, Cubberley writes as follows:

"These associations of scholars, or teachers, or both, 'born to the need of companionship which men who cultivate their intelligence feel,' sought to perform the same functions for those who studied and taught that the merchant and craft guilds were performing for their members. The ruling idea was association for discussion and study; the obtaining of corporate rights and responsibilities; and the organization of a system of apprenticeship, based on study and developing through journeymen into mastership, as attested by an examination and the license to teach." 10

Tests still depended solely on the teacher and his teaching and might vary endlessly. Oral tests gave way to written tests and through the centuries tended increasingly to be the yardstick by which progress was measured. By the 19th century formal examinations in all branches of study were well established throughout Europe. As our concern is

9
Ellwood P. Cubberley. The History of Education.
New York: Houghton Mifflin Company, 1920, p. 93-4.

10
ibid., p. 217.

primarily with testing in school learning, we can limit our observations to this field. England is taken as our example because her educational structure, through the American Colonies, has been the foundation of our system and hence is directly related to the investigation.

In England, as elsewhere, the first schools were Church schools and, as such, were under the jurisdiction of the Church. Regulations and standards depended entirely upon the Church concerned. Later, organized education received grants from church-school societies and the government, as a result, demanded the right to inspect schools. Usually the government appointed clergymen as school inspectors because of the power of the Church and because they were the best educated men of that day. However, there was friction between Church and State and finally, in 1870, State Schools were established in which we find local school boards with certain definite duties and powers. Rigidly controlled examinations began to dominate the educational system and, as shown by the following quotation, shackled it severely:

"Secondary education had its outlook narrowed and its aims warped by the necessity to prepare for certain competitive examinations which were supposed to set upon the schools the hall-mark of success. So keenly was the pressure of examinations felt that in November 1888 there appeared in the 'Nineteenth Century' an article protesting in the strongest terms against our 'sacrifice of education to examination', an article signed by 400 of the most eminent men and women of the day." 11

11
P.B. Ballard. The New Examiner. London: University of London Press, 1929, p.17.

In North America we find transplanted systems of European Education. A public school system, founded on that of England, was established in New England by the 17th century. Each State tended to vary somewhat depending upon the racial background. As the examination system was prevalent throughout all European schools, it was adapted to American schools.

"It was customary in the early days for the school committees in Massachusetts to give oral examinations in the schools under their control. By 1845 the enrollments had become so large in Boston that the committee could no longer devote the time required for anything more than the casual examination of each pupil with an oral quiz. To meet this situation the uniform written examination was adopted. The results were so gratifying that Horace Mann wrote his enthusiastic defense of written examinations. . . ." 12

In Massachusetts we have the first free schools, and this state has always been one of the fore-runners in educational progress. Later, about 1850, Horace Mann gave a fine criticism of examinations in Massachusetts and the following quotation shows that he was well ahead of his day in educational outlook:

"Horace Mann, for example, about 100 years ago, had a remarkable conception both of the importance of examinations and of the limitations of the forms then in existence. His penetrating analysis of the weakness of the oral examinations then in vogue, and of the superiority of written examinations, could hardly be improved upon by the modern specialist in measurement. Mann showed clearly the points where the oral examinations were lacking, in the technical language of to-day, in validity, reliability and usability." 13

12

C.C. Ross. op. cit., p. 51.

13

ibid., p. 29.

The examination system became generally and widely accepted until it acquired an importance that threatened to dominate the process of learning.

"Those examinations which have been given at the end of a marking period or a school term—commonly called final examinations—were particularly over-emphasized in the past. In many schools the pupil's promotion or failure depended entirely on his performance on the final examinations. This alone magnified the significance of these examinations in the minds of teachers, pupils, and parents, all out of proportion to their actual educational value. The way in which the examinations were approached and the manner in which they were administered added to the dread that surrounded this phase of the child's educational experience. The setting aside of special examination days, the comprehensive reviews, the warnings of the teacher, and the formal examination atmosphere — all these were part of a pattern which led pupils to approach final examinations with varying degrees of concern." 14

The marking system was a direct development of the examination system. In an effort to measure pupils' ability and educational progress, teachers endeavored to work out a method of marks to represent more accurately the standard reached by the pupils. At first the chief basis for marks was absorption of content, measured by some standard which was supposedly applied uniformly to all pupils regardless of their abilities. This system of marking placed the child of low intelligence in direct competition for marks with his more able class mates. Obviously, he was at a hopeless disadvantage from the start. Under that system the hard-working child of low intelligence might fail again and again, while the brilliant pupil with little effort might receive the

highest marks. Pupils who were a few points below the passing mark were branded as failures in school records. Students were passed who reached the "passing mark". Investigations have shown that many pupils who failed school grades have been successful in university or in business life. Everyone knows of "poor students" in school who were a great success in life. It has been established that many scores were an inaccurate indication of a pupil's ability and potential accomplishment.

Period of First Scientific Measurement in Education

As the natural sciences made rapid progress in exact measurement, leading educationalists and research men turned more and more to measurement in the educational field. Many thought that if education was to be scientific it must be directed by psychology. In 1885, G.F. Hall published an outstanding work based on psychology called "Adolescence". It had a profound influence and child-study acquired widespread importance. One of the first reactions to this new phase was the evolution of better methods of investigating educational measurement.

Another pioneer in psychological study was J. McKeen Cattell who made efforts to develop tests of mental ability by means of which the differences between individuals could be compared and the changes through which individuals pass in the course of development measured.

"More than anyone else, Cattell was responsible for giving to American psychology its practical bent, for with him the practical took precedence over the philosophical. As early as 1885 he began to publish important articles on reaction times and individual differences. It is Cattell who in 1890 suggested the term 'mental tests', which was to become a sort of trade-mark for the whole measurement movement." 15

His first results were comparatively unsatisfactory but were later perfected and are now standard instruments of measurement in psychology and education. Cattell and other American educators of the same period were actually carrying out and expanding somewhat the achievements of Wundt, of Leipzig, Germany, who was the first man to establish a laboratory in psychology.

"Measuring instruments also make it possible for the worker to resort to experimental methods to learn definitely whether materials and methods are effective. This is true in the field of teaching as in other fields. Without specific aims the teacher cannot plan his work effectively. He cannot know, except in an indefinite way, what he is to do." 16

Prior to 1894 little progress was made in combatting the formal examinations of subject matter only. It was in this year that J.M. Rice conducted his spelling study in many schools and later his arithmetic investigation. He published an excellent article showing the necessity for scientific testing of school results. He is considered the real inventor of the comparative test although educational leaders frowned on his work and it required others to perfect it. The work of these men in America closely parallels

15

C.C. Ross. op. cit., p.36.

16

W.T. Gruhn-H.R. Douglass. op. cit., p.10.

advances made by two well known educationalists outside this continent.

Galton, an Englishman, around 1883 produced his statistical analysis for the measurement of intelligence. He was first in England to study intensively the problem of individual differences in psychology. Across the English Channel, Binet, a Frenchman, was attempting to work out a method of measuring intelligence. A man of great imaginative powers, by the method of repeated trial and error and by copying freely from others, he finally in 1905 produced the first scale for the measurement of intelligence. It is true that it was imperfect, but it served as a base for all future measurement scales. Goddard translated Binet's 1905 scale and adopted it in America.

Following closely upon the Binet Scale, numerous tests were devised measuring the rate and degree of accuracy in subject matter among pupils of different grades. It is to the genius of Thorndike that we owe the next step. He devised a scale unit for the measurement of educational achievement. No other person has studied the measurement movement in so many ways or has contributed so much to it. His work was so extensive and so effective that he is considered the "father" of the movement.

"Stimulated by the work of Dr. Rice, Professor Thorndike began to experiment with tests and scales. Stone, Hillegas, Buckingham, Trabue, Curtis, and many others were stimulated by his teaching, with the result that within the decade following Rice's introduction of the measurement idea many significant contributions were made to educational measurement." 17

One more man deserves a word of mention among those associated with scientific educational measurement. Lewis M. Terman of Stanford University produced his Stanford-Binet Scale in 1916 with his manual "The Measurement of Intelligence". For two decades it remained the best individual intelligence test.

Other men were instrumental in extending the science of measurement in education during this period but the former ones made outstanding contributions. Acceptance of these advances, however, was not immediate. Education is the concern of all society, and opposition from nearly every group of organized life builds up against any new system that may react against its interests. Revolution, accordingly, or even progress, in education is necessarily slow. Even in the ranks of educationalists themselves, opposition is often strong and teachers are slow to accept new theories. Teachers for a time expressed curiosity in the new tests and scales but felt at first that they belonged to experts only. Credit for popularizing the rise of these new tests belongs to Courtis who, in 1911-12, successfully disseminated interest in them through a survey.

What were some of the main factors accelerating this measurement idea?

(1) Educators discovered how inadequate were actual measurements in schools. Many surveys in schools and colleges revealed that school marks were too subjective. Evidence was accumulated to show that teacher's marks varied as to time

and as to subject. In a Wisconsin high school survey C.E.

Hulten reports:

"that 15 teachers who gave passing marks the first time would have failed the pupil the second time the paper was marked and that 11 teachers who gave failing marks the first time would have passed the pupil the second time." 18

Marks were too frequently a reflection of the personality of the teacher rather than a true indication of the pupil's ability. Thorndike and other educators were shocked by the results. Even lay-teachers saw the need for more accurate measures.

(2) Dissatisfaction with existing marks assigned by teachers. Variations in teachers' marks caused great concern to professors of education in laboratory schools.

"Why do teachers differ so much in estimating the worth of a given product and in the distribution of marks to groups of pupils? Four possible factors may be mentioned: (1) Differences in the standard of severity or leniency of different teachers; (2) Differences in the standards of severity or leniency in different schools; (3) Differences in credit or penalty assigned by different teachers to any given fact or error in a piece of work; (4) Minuteness of the discrimination between successive steps of merit or quality in a given scale of qualities." 19

(3) Research Bureaux in large educational systems were extended more effectively and examined the use of educational tests. All associated people were trained in their use and, realizing their educational importance, they carried the new testing techniques back to the classrooms.

18

C.C. Ross. op. cit., p.47.

19

H.A. Greene-A.N. Jorgensen-J.R. Gerberich. Measurement and Evaluation in the Secondary School. New York: Long. 1943, p.136.

The recognition of the value of standard tests for diagnostic purposes has done much to popularize their use and to transfer them from the hands of specialists to teachers. The new-type tests are more objective, they can be marked easily, they permit wider sampling of the subject, and they effect a more reliable mark than the old essay-type examinations. Of course these tests were immediately popular, as many teachers believed the ordeal of final examinations would be over. These tests, however, have limitations as well. If poorly constructed, they tend to encourage the memorization of facts, they fail to develop skill in solution of problems, and they do not allow pupils to form their own opinions Mursell states that:

1. "they do not reveal a person's capacity for complex and sustained learnings.
2. our tests cannot directly reveal capacity for disentangling concepts from complex situations.
3. our tests cannot directly or indirectly reveal capacity for consistent and considered choice between possible courses of action.
4. our tests cannot directly reveal capacity for dealing sensibly and wisely with practical problems.
5. our tests cannot reveal directly a person's capacity for controlled and effective methods of work.
6. our tests cannot directly indicate the depth, strength and subtlety of a person's appreciative reactions in ethical, social, or aesthetic matters.
7. above all, our tests cannot even begin directly to reveal capacity for producing original ideas and construction — for initiative." 20

"The other side of the picture is that despite limitations which every judicious student of the subject is bound to recognize, the modern testing movement has achieved great and indubitable successes, both practical and theoretical." 21

The following quotation summarizes the evolution of testing systems and the progress towards efficacy of tests as measurements in evaluating educational progress:

"Teachers long have measured the results of their teaching efforts. However, it is only relatively recently that any degree of accuracy has been injected into their methods of measurement. For many years the teacher's estimate was accepted as the sole measure of a pupil's ability or accomplishment. Studies of the reliability of such methods gradually began to cast a doubt on their accuracy. Accordingly, interested teachers and research workers began to seek for more dependable measures. This movement was fostered by the so-called 'survey' movement among educators which appeared at about the same time. Possibly the survey movement itself was a product of the same spirit of unrest and dissatisfaction with educational methods which brought into being the measurement movement. The survey movement left in its wake a perfectly logical result—the establishment of many centres of interest in the more exact evaluation of the results of educational practices. These later appeared as bureaus of educational measurement and research, a few of which are still functioning. Very distinctive service has been rendered by these agencies through their work in construction, standardization, and critical evaluation of educational measuring instruments." 22

21
ibid., p.16.

22
H.A.Greene-A.N.Jorgensen, op.cit., p.572.

Progress since 1918

Thorndike in 1918, published one of the most influential papers of this period on educational measurement.

"The paper began with the well known dictum: 'Whatever exists at all exists in some amount', and ended with this note of satisfaction: 'Of the gains made in the past decade, we may well be proud.'" 23

In this paper Thorndike showed the gains of the past decade and predicated much improvement in the immediate future. The place of these tests was finally established in the field of education and their use began to spread into other lines of work. A system of testing had been initiated in army life during the First World War and was considered of value. Buckingham 24 states that in 1919 "test-making passed from an amateur to a professional basis". The general picture is that, as in schools teachers had accepted test to improve marking, so the general public adopted unconditionally the test-result method. In fact we enter a decade or so of complete confidence in all tests. Many were good but inevitably some were poor. One point is clear: Monroe 25 in his "Ten Years of Educational Research, 1918-27" notes "that the pioneer state of educational research is passed."

By 1927 the survey type of tests had given way to specific and diagnostic tests. Their purpose was to show a

23

C.C. Ross, op. cit., p. 49.

24

R.B. Buckingham. "Our First Twenty-five Years."
Proceedings of the National Education Association, 1941, p. 354.

25

Walter S. Monroe. "Ten Years of Educational Research?
1918-27. Bureau of Educational Research Bulletin, No. 42.
Urbana: University of Illinois, 1928.

pupil's weak and strong points and they were a decided improvement on the early tests, which were too general. The rapid development of high-school tests for all subjects followed. These were a great factor towards better marking as objective tests were cheap and so were widely used. About this time some educators began to doubt the results of the tests and a critical attitude developed towards them. As the field of measurement and experiment was extended into new areas, the critical force increased and, as it was supported by educational publications of the day, it resulted in more intensive diagnostic work. In view of this research, more attention was given to reliability and validity of tests. The final outcome of this later period was that people realized that standard tests did not meet all the required needs of measurement and techniques must be improved for truer evaluation. This critical attitude towards tests is a strong progressive factor because it requires of teachers, not that they should abandon objective tests, but that they should analyze the test scores.

"Hildreth points out some beneficial results of this change: A more critical attitude towards intelligence measurement, as the outcome of continued experimentation, has resulted in more authoritative research findings, more sensible and intelligent interpretation of data. This attitude has shown itself with respect to achievement tests and personality measurements as well. The result has been not so much the curtailment of the use of tests, as their more critical use and the more cautious interpretation of test scores." 26

The gradual progress in scientific measurement with its resulting standardized and diagnostic tests has led steadily to greater validity in teachers' marks and has reacted against the general tendency to discredit the examination system. The present study may be considered, therefore, as an indication of the value of teachers' marks in the high schools of Manitoba for the purpose of promotion from Grade XI.

CHAPTER III

A STUDY OF STUDENT MARKS IN SIX COLLEGIATES FOR 1949

The investigation was begun in Gordon Bell High School by obtaining Department of Education June marks for non-exempted students for 1949. The writer obtained from school records marks for students listed for non-exemption the previous May. These marks were recorded on Tally Sheets (see Appendix A). The name of the student was replaced by a number in the final analysis of totals. The results, recorded in Table II read as follows: Subject, Number of Pupils, Total (whether positive or negative), per Pupil Average, either positive or negative, and, lastly, the Number of Papers having the same mark as that of the Department of Education. These results are recorded in Tables II to VII inclusive, of this chapter.

^ Comparative study of the marks of Gordon Bell Collegiate is presented in Table II.

TABLE II

SUMMARY OF GORDON BELL EXAMINATION RESULTS COMPARED
WITH DEPARTMENT OF EDUCATION MARKS FOR THE YEAR 1949

Subject	Number of Papers	Total Plus or Minus	Per Pupil Average Plus or Minus	Number of Papers Equal to Dept.
Lit.	68	-82	-1.16	3
Comp.	70	245	3.5	2
Hist.	64	325	5.01	1
Geom.	67	-1232	-18.39	0
Alg.	65	-519	-7.98	0
Chem.	59	-592	-10.	1
Phys.	21	-170	-8.1	1
Fr.	50	-386	-7.6	1
Lat.	13	-119	-9.1	0

This table represents the measure of difference in teacher and Provincial marks for each of the nine subjects. The subject in first position is Literature with -1.16 average and it means that each pupil was marked 1.16 points below the mark obtained in the Provincial Examination held in June. Composition is next with 3.5 average, followed by History. These are the only subjects over-scored. Algebra and French have the same average followed closely by Physics while ⁱⁿ Latin the range is a little wider. Chemistry with -10. has the second highest under-valuation, with Geometry, as the highest, averaging -18.39. The general results presented here indicate, for 1949 at least, that teachers under-estimated the ability of the students.

It was considered advisable to carry the study to another Winnipeg High School to ascertain the degree of similarity or variation from school to school within the City of Winnipeg and to establish the reliability of the sampling. Marks for "non-recommended" pupils from Kelvin High School were secured for this purpose. The results from Kelvin are presented in Table III.

TABLE III
SUMMARY OF KELVIN EXAMINATION RESULTS COMPARED
WITH DEPARTMENT OF EDUCATION MARKS FOR THE YEAR 1949

Subject	Number of Papers	Total Plus or Minus	Per Pupil Average Plus or Minus	Number of Papers Equal to Department
Lit.	69	101	1.46	4
Comp.	70	486	6.94	0
Hist.	68	352	5.17	1
Geom.	65	-995	-15.31	1
Alg.	65	-414	-6.37	0
Chem.	70	137	1.95	3
Phys.	24	-60	-2.5	0
Fr.	52	344	6.61	0
Lat.	8	156	19.5	0

The Kelvin school returns have many results in common with Gordon Bell for 1949. The two schools have almost the same number of students. Literature again has the best average and also has four papers with equal marks in Collegiate and Departmental scores. In fact Literature and History are very

close in both schools. Science marks bear a similar relation to those of the Department. Chemistry is over-estimated, Physics under-estimated. Geometry has the greatest under-^{ation}valuing, -15.31, and Algebra has -6.37. In Mathematics the averages are similar in the two schools. In the Language averages the situation is exactly the reverse of that of the first Winnipeg school. Both Languages are decidedly over-estimated, while in Gordon Bell, they are as much under-estimated.

Each school has a total of nine sets of equal marks.

A comparison of marks for the West Kildonan (Centennial) Collegiate is reported in Table IV.

TABLE IV

SUMMARY OF WEST KILDONAN RESULTS COMPARED WITH
DEPARTMENT OF EDUCATION MARKS FOR THE YEAR 1949

Subject	Number of Papers	Total Plus or Minus	Per Pupil Average Plus or Minus	Number of Papers Equal to Department
Lit.	15	14	.93	0
Comp.	16	267	16.75	0
Hist.	16	196	12.5	0
Geom.	15	9	.66	1
Alg.	15	1	.06	0
Chem.	15	9	.66	0
Phys.	15	159	10.6	1
Fr.	9	1	.11	1
Lat.	4	-5	-1.25	2

This Collegiate has the best returns of all schools examined and presents many close averages, namely, Literature with .93, Geometry .6, Algebra .06, Chemistry .6, and French

.11 per paper. Latin with -1.25 average is exceptionally high for that subject and is outstanding in the respect that there are two sets of equal marks out of four papers.

In these results Algebra and Geometry have almost the Departmental average. The two previous schools show considerable under-estimation in both Mathematics. Literature has a close average but Composition is decidedly over-estimated. These English subjects have no papers with the same marks as those of the Department. History is also over-marked.

With the exception of Composition, Physics and History West Kildonan has all averages close to the Department standard.

The analysis of marks for the Dauphin Collegiate are reported in Table V.

TABLE V
SUMMARY OF DAUPHIN EXAMINATION RESULTS COMPARED
WITH DEPARTMENT OF EDUCATION MARKS FOR THE YEAR 1949

Subject	Number of Papers	Total Plus or Minus	Per Pupil Average Plus or Minus	Number of Paper Equal to Department
Lit.	34	-36	-1.06	0
Comp.	34	-54	-1.58	0
Hist.	34	35	1.03	0
Geom.	32	-70	-2.18	0
Alg.	30	-60	-2.0	1
Chem.	34	-111	-3.27	0
Phys.	27	-18	-.66	1
Fr.	22	8	.36	0
Lat.	11	-33	-3.0	0

These results are the best so far because there is no significant over-estimation or under-estimation. It is true that seven subjects are under-scored but not to any degree. Averages are close throughout although there are only two equal marks.

The best average is French with .36, closely followed by Physics. The two English subjects have a near zero average and present less spread than in other schools. History also presents a close average. Algebra and Geometry show a slight deviation from the Departmental standard. Latin has a close average, -3, and is followed by Chemistry with the widest variation, -3.27.

The analysis of the Morden (Maple Leaf) Collegiate returns are given in Table VI.

TABLE VI
SUMMARY OF MORDEN EXAMINATION RESULTS COMPARED
WITH DEPARTMENT OF EDUCATION MARKS FOR THE YEAR 1949

Subject	Number of Papers	Total Plus or Minus	Per Pupil Average Plus or Minus	Number of Papers Equal to Department
Lit.	9	-57	-6.3	0
Comp.	9	152	16.89	0
Hist.	10	15	1.5	0
Geom.	10	74	7.4	0
Alg.	8	97	12.12	0
Chem.	8	-80	-10.	0
Pyhs.	7	-51	-7.28	0
Fr.	7	52	7.43	1
Lat.	4	24	6.0	1

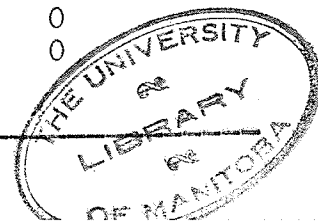
These averages represent more diversified returns than in other schools. The best average, 1.5, is in History. Latin, over-scored by six points, comes next. Literature is underestimated by the same amount. Geometry and French have the same average, with Physics under-valued to the same degree. Chemistry is considerably underestimated, and Algebra and Composition have the widest deviation.

Languages are over-estimated as in Kelvin. Only in Languages are there any sets of marks the same in both school and Departmental results. Chemistry and Physics are similar to the two Winnipeg Collegiates— in fact Chemistry is the same as in Gordon Bell school. In Mathematics we find much over-estimation, particularly in Algebra. This is the first Collegiate with such significant over-scoring.

The last Collegiate studied was Portage la Prairie. Table VII contains its summary.

TABLE VII
SUMMARY OF PORTAGE LA PRAIRIE EXAMINATION RESULTS COMPARED
WITH DEPARTMENT OF EDUCATION MARKS FOR THE YEAR 1949

Subject	Number of Papers	Total Plus or minus	Per Pupil Average Plus or Minus	Number of Papers Equal to Department
Lit.	18	39	2.16	1
Comp.	18	84	4.66	0
Hist.	18	153	8.5	1
Geom.	18	-177	-9.83	0
Alg.	18	-138	-7.66	0
Chem.	18	193	10.73	0
Phys.	14	168	12.0	0
Fr.	11	-129	-11.72	0
Lat.	7	-60	-8.57	0



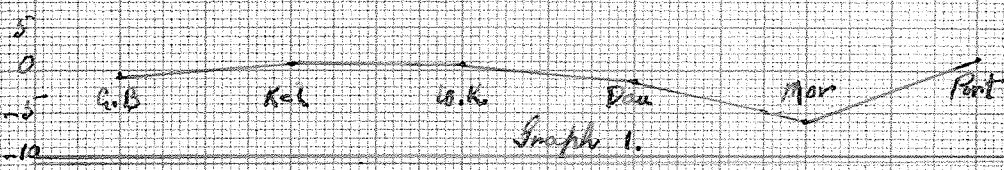
The pattern of averages in this Collegiate presents one unusual feature. Literature and Composition are both positive, Mathematics are both negative, Sciences are both positive, and Languages are both negative. In other words, each group is entirely positive or entirely negative.

In the English section, the results are similar to those of Kelvin and Gordon Bell, although Literature is over-valued. It will be noted that Literature has the closest average and has one paper equal to that of the Department of Education. History is over-scored more than in any other school except West Kildonan. Geometry and Algebra are both under the standard as in Dauphin, Kelvin and Gordon Bell. Languages bear out the under-valuation of Gordon Bell. Sciences are decidedly over-estimated, a situation opposite to the results in the majority of the other Collegiates.

Summary of Subject Matter Variation, Study I

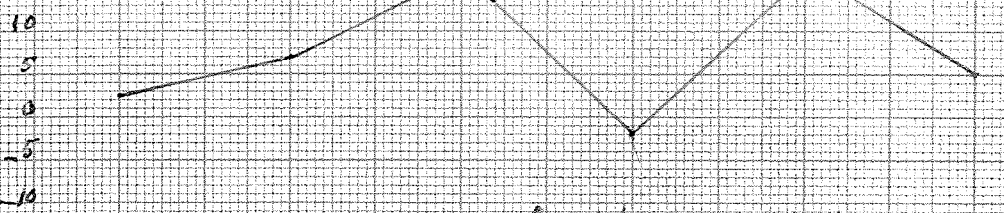
The remainder of the study on variation in examination marks for 1949 consists of graphical presentations of variation. The charts are made for each subject individually and then in composite graph form. A deviation from the Departmental average, which is represented by zero, has the same value whether plus or minus. In other words a +4 rating represents the same distance from zero as does a -4. Graphs are presented hereafter for each subject and on page 45 a composite graph is given for all subjects.

10 Lit.



Graph 1.

15 Comp.



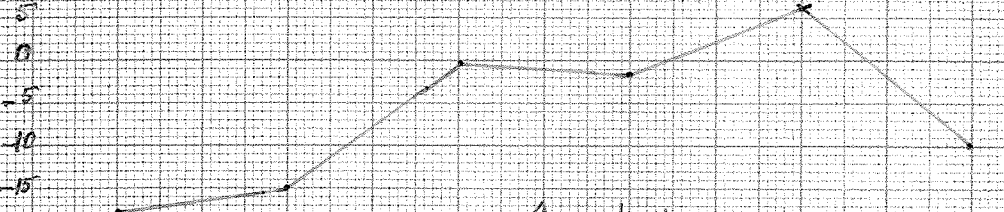
Graph 2.

15 Hist.



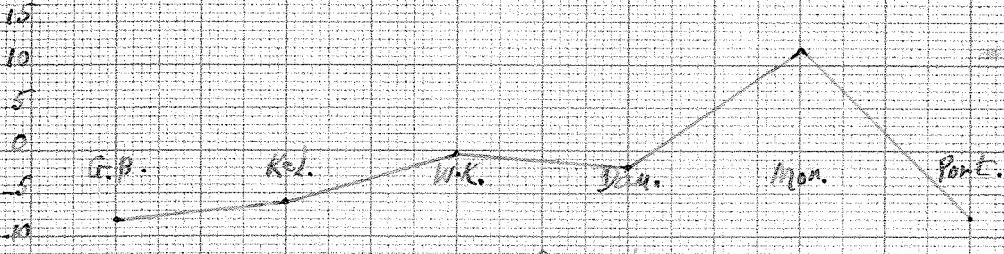
Graph 3.

10 Geom.



Graph 4.

12 Alg.



Graph 5.

Figure 1.

Variation in Subject Averages for 1949.

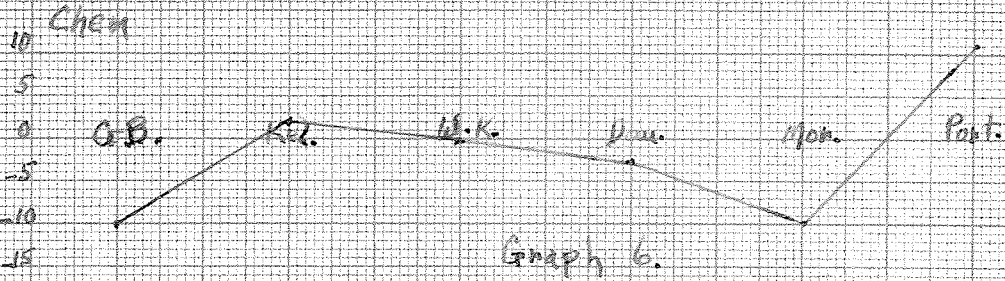
IF SHEET IS READ THIS WAY (HORIZONTALLY), THIS MUST BE TOP.
IF SHEET IS READ THE OTHER WAY (VERTICALLY), THIS MUST BE LEFT-HAND SIDE.

THIS MARGIN RESERVED FOR BINDING.

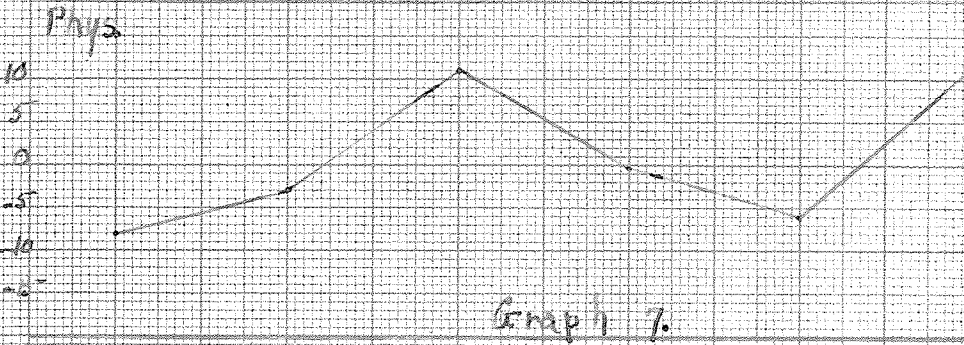
THIS MARGIN RESERVED FOR BINDING.

IF SHEET IS READ THIS WAY (HORIZONTALLY), THIS MUST BE TOP.

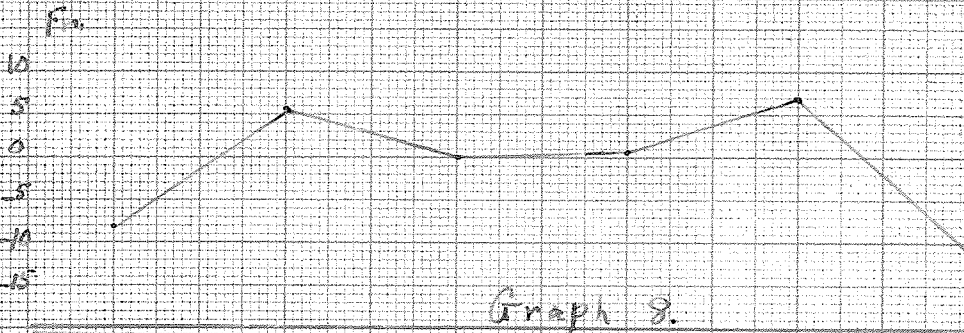
IF SHEET IS READ THE OTHER WAY (VERTICALLY), THIS MUST BE LEFT-HAND SIDE.



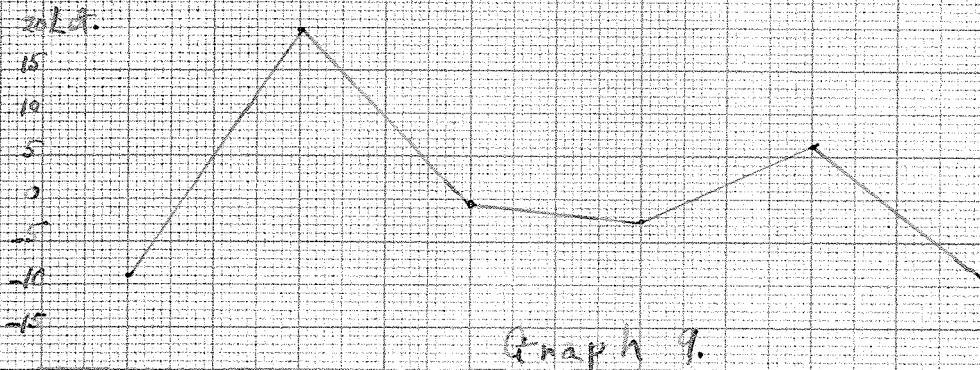
Graph 6.



Graph 7.



Graph 8.



Graph 9.

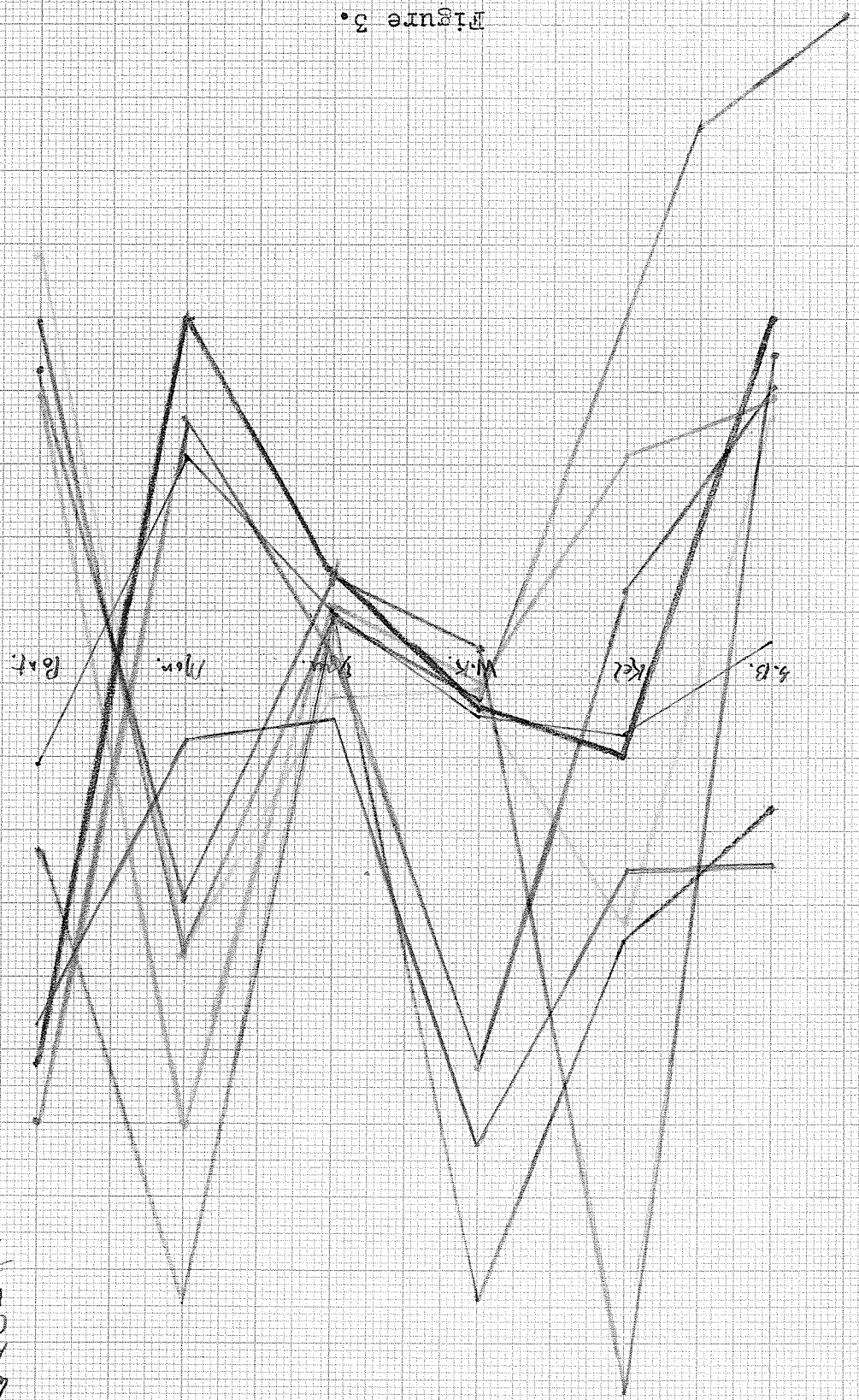
Figure 2.

Variation in Subject Averages for 1949.

20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0
-1
-2
-3
-4
-5
-6
-7
-8
-9
-10
-11
-12
-13
-14
-15
-16
-17
-18
-19
-20

Composite Graph on Variation for 1949.

Figure 3.



- Legend
- Lot. - Blue Ink
 - Comp. - Dark Blue
 - Hof. - Pink
 - Green - Brown
 - Red - Light Blue
 - Cham. - Red
 - Play. - Green
 - Pr. - Yellow
 - Lot. - Indefinite

The previous graphs for each subject examined in 1949 present their variations. These, in turn, are co-ordinated in a composite graph. It is not the intention of the writer to repeat individual subject graphs for the years 1950 and 1951, but rather to use the composite graph only. In Figures 1 and 2, variations are shown. The writer considers from +5 to -5 a reasonable fluctuation from the Department of Education standard. This is, of necessity, an arbitrary area of measurement but will serve adequately to illustrate variation. Figure 3 shows that all subjects for one Collegiate are inside the designated range and six subjects for another school are also within this area. In the remaining schools are to be found a varying number of subjects beyond this range. In the latter schools it is usually a few subjects that have caused this wide variation.

As shown in graphs 1, 2, 3, 8 and 9, slight variations exist in Literature, Composition, History, French and Latin. In the remaining graphs, curves are skewed because of extreme variation found in some cases in occasional subjects. History is over-estimated in every Collegiate examination, and Composition in five out of six.

The last part of the investigation for this year is a comparison of the means obtained from Departmental marks and those of the Collegiates. Table VIII contains these data.

TABLE VIII

RANGE IN MEANS OF SCHOOL MARKS COMPARED WITH
THE RANGE OF DEPARTMENT OF EDUCATION MARKS, 1949

	G.Bell	Kelvin	W.Kildonan	Dauphin	Morden	Portage	Range
Lit. Mean	60	55	65	56	54	61	54-65
Dept. "	62	54	62	60	59	60	54-62
Comp. "	62	58	62	54	64	63	54-64
Dept. "	57	51	38	53	51	58	38-58
Hist. "	61	55	57	60	64	69	55-69
Dept. "	58	51	50	61	56	63	50-63
Geom. "	63	45	50	58	77	58	45-77
Dept. "	80	56	50	63	61	69	50-80
Alg. "	65	50	51	69	76	58	50-76
Dept. "	72	57	55	73	58	71	55-73
Chem. "	54	53	43	62	36	70	36-70
Dept. "	65	51	40	60	52	62	40-65
Phys. "	59	53	50	70	42	67	42-70
Dept. "	64	57	40	71	51	58	40-71
Fr. "	58	61	50	55	50	55	50-61
Dept. "	64	54	53	56	39	66	39-66
Lat. "	50	60	55	60	50	54	50-60
Dept. "	57	50	55	64	43	64	43-64

On comparing the means of the subjects as listed in this table, the former variations of the subjects, in the main, are upheld. We find the range is lowest in Literature when both Department of Education and Collegiate returns are considered. This means that the results are best in this subject. History follows, with the range almost equal in both Departmental and Collegiate marks. Composition, Latin and French form a satisfactory group.

The lowest group, in keeping with the previous analysis, consists of Chemistry, Physics, Algebra and Geometry in which the spread from the means is wide in both Collegiate and Departmental results.

In conclusion, some significant results may be re-stated in brief for the study of examinations made for 1949.

1. In two Collegiates the fluctuation is slight from Department standards. In the other Collegiates wider variation has been caused, chiefly by a few subjects in each case.
2. The variation in subject matter scores is not large for Literature, History, Composition, Latin and French.
3. There is significant variation in the case of Chemistry, Physics, Algebra and Geometry.
4. History is over-estimated in all six schools, and Composition in five of the six Collegiates.

CHAPTER IV

COMPARATIVE STUDY FOR THE YEAR 1950

The next year studied in the investigation is 1950. The returns for that year are herewith submitted for subsequent analysis. Table IX presents the Gordon Bell results for 1950.

TABLE IX

SUMMARY OF GORDON BELL EXAMINATION RESULTS COMPARED WITH DEPARTMENT OF EDUCATION MARKS FOR THE YEAR 1950

Subject	Number of Papers	Total Plus or Minus	Per Pupil Average Plus or Minus	Number of Papers Equal to Department
Lit.	51	-175	-3.43	3
Comp.	52	312	6.0	0
Hist.	54	360	6.74	1
Geom.	49	-482	-9.83	0
Alg.	3	-27	-9.0	0
Chem.	17	-315	-18.52	0
Phys.	12	11	.91	1
Fr.	3	-42	-14.0	0
Lat.	0			

Physics shows the least variation, being only .91 above Departmental average. Literature is under-valued -3.43 while Composition and History are almost equal, an over-valuation of 6 and 6.74 respectively. Geometry and Algebra have approximately an equal negative rating. French with few papers is under-valued 14 points. Chemistry has the greatest variation with -18.52 average.

It was considered advisable to select the middle year, 1950, for a study of pupil variation. The mark of each pupil in any subject the average for which was outside the ± 5 to -5 range, was examined with a view to determining the number of individual papers in that subject beyond a ± 10 to -10 variation from the Departmental standard. To comprehend the division more easily, two separate columns are made for the papers over-valued with the division ± 10 to ± 20 in one column, and over ± 20 in the other. The same recording was made for the under-valued papers. This has been done for each Collegiate and is reported for Gordon Bell in Table X. Deviation of individual pupil marks is indicated by the number of papers recorded in the columns headed Above Department Average and Below Department Average. For example, in Gordon Bell, Chemistry has an average of -18.52 (Table IX), and the number of pupils in the under-valued column is 12, or 12 papers out of a total of 17.

TABLE X

PUPIL VARIATION IN WIDE RANGE SUBJECTS IN GORDON BELL, 1950

Subject	Number of Papers	Number of Pupils Above Dept. Average		Number of Pupils Below Dept. Aver.	
		± 10 to ± 20	± 20 --	-10 to -20	-20 --
Comp.	52	13	6	5	0
Hist.	54	15	7	6	1
Geom.	49	4	0	10	13
Alg.	3	0	0	1	1
Chem.	17	0	0	6	6
Fr.	3	0	0	1	1

Geometry has 23 out of 49 pupils below the -10 range and Chemistry has 12 out of 17 papers so defined. Composition and History are over-valued. The school marks in Geometry tend to be severe. On the other hand marks in Composition and History would indicate leniency. These variations support the findings of Table IX.

The results for Kelvin Collegiate are reposted in Table XI.

TABLE XI
SUMMARY OF KELVIN EXAMINATION RESULTS COMPARED
WITH DEPARTMENT OF EDUCATION MARKS FOR THE YEAR 1950

Subject	Number of Papers	Total Plus or Minus	Per Pupil Average Plus or Minus	Number of Papers Equal to Department
Lit.	43	-227	-5.27	1
Comp.	42	-3	-.07	0
Hist.	46	131	2.84	1
Geom.	35	-93	-2.65	0
Alg.	21	197	9.38	0
Chem.	21	-67	-3.19	0
Phys.	3	2	.66	0
Fr.	17	126	7.41	1
Lat.	11	56	5.09	1

Composition shows the least variation, being only -.07. Physics, with three papers, is .66. Subjects varying significantly are Literature, Algebra, French and Latin. History and Geometry average 2.84 above and -2.65 below respectively. Chemistry has -3.19 under-estimation. French shows an average of 7.41 while Algebra is slightly wider in its variation.