1954

Evaluation of teacher-made elementary algebra tests in the secondary schools of Montana

Duane Leonard Fredricks
The University of Montana

AN EVALUATION OF TEACHER-MADE ELEMENTARY ALGEBRA TESTS IN THE SECONDARY SCHOOLS OF MONTANA

by

DUANE L. FREDRICKS

B. S. in Education, University of North Dakota, 1952

Presented in partial fulfillment of the
requirements for the degree of
Master of Arts

MONTANA STATE UNIVERSITY

1954

Approved by:

Chairman, Board of Examiners

Dean, Graduate School

Aug 13 1954

Date

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

A STATEMENT OF THE PROBLEM

The use of the teacher-made test has been the principal means for determining the grade of the pupil. The grade in turn has been used as a basis for deciding whether or not the pupil is ready for promotion or graduation. Although the concept of the grade may be overemphasized in the educational system, it nevertheless exercises a considerable influence over the pupil. In view of these considerations, every effort should be made to insure that the grade be as accurate as possible.

The teacher-made test may be divided into two main categories, the objective and the subjective. The latter is better known as the essay-type of test. The objective test has come into vogue in recent years and is now widely used as an evaluating instrument. The essay or subjective-type test, while still being employed in the school, has been accused of lacking in validity because it does not measure what it is supposed to measure and of lacking in reliability because, since it is dependent on the opinion of the teacher, it often fails to measure consistently what it does measure. The objective-type of examination was

introduced in an attempt to remedy many of the evils of the essay examination. In various school subjects, however, such as history, English, and even the sciences, the teacher on an objective examination could be predisposed toward certain answers in which there would be room for argument. Many believe that this difficulty would not exist in any mathematics examination, since such an exact subject would not permit variation in answers.

This study dealt with one branch of mathematics, namely, elementary algebra. It was limited to the teachers of this subject in the secondary schools of Montana.

One of the aspects of the study was to examine and evaluate the current testing procedures of the teachers of elementary algebra. Two types of elementary algebra tests were taken into consideration. One type was the major test given at the end of a report period, for example, yearly, semester, or six weeks; the other was the minor test given daily or at less frequent intervals. Information concerning the preparation, administration, and evaluation procedures of these two types of tests was obtained by means of a questionnaire sent out to a sample of Montana teachers of algebra.

One part of the questionnaire was devoted to the determination of the degree of agreement among teachers on the scoring of two elementary algebra problems. One

problem concerned the solution of a fractional equation; the other was a statement problem giving a fractional equation to solve when set up. Errors were deliberately introduced into the solutions of each problem.

The other phase of the study was concerned with the teacher-made test itself. Each teacher was requested to send in a copy of a major elementary algebra test which had been given recently to his class. These tests were evaluated in terms of their validity and reliability.

IMPORTANCE OF THE PROBLEM

The problem is important because the test is used to determine the grade of the pupil. Such grades are ordinarily averaged to form a final grade. The final grade dictates the future of the pupil in that subject or line of endeavor. Some feel the mark is overemphasized and that the grading system should be eliminated as a result. Others feel that drastic changes such as reducing the marking system to a pass or fail basis should be made. Whether or not these changes should be made is debatable. For the present, however, the marking system is still in use.

If the marking system is to exercise a considerable influence over the pupil, the accuracy and the precision of the evaluating instruments that determine such marks should be of paramount consideration. The validity of the test is a measure of the accuracy of the results and the reliability a measure of how precise the results are. The

problem is important from the aspect of validity and reliability.

If every mathematics problem has a definite answer, marking in an objective manner should be simple. The more objectively a problem can be marked, the more agreement there should be among those who grade it. If there is disagreement in the evaluation of the pupil's work on the problem, it is probably because of other factors entering into the evaluation. These factors may contribute to the subjectivity of the marking procedure. It is important that these factors be studied as to their actual relation to a good evaluation procedure.

The procedures of the teacher in the preparation, administration, and evaluation of the test determine whether or not the test will be a good measuring instrument. Therefore, a study of these procedures appears to have merit.

SOURCES OF INFORMATION

As indicated above, part of the material for the study was compiled from a questionnaire sent to elementary algebra instructors in various schools in the state of Montana. Alternate schools were chosen from those listed in a pamphlet[1] issued by the State Department of Public Instruction of Montana. Each alternate school, giving a total of 100, was separated according to the three classes

---

[1]Montana Educational Directory, State Department of Public Instruction, Helena, Montana, 1952-1953.

of districts, the county high schools being placed in a separate classification.

The first class district is defined as follows:[2]

> A first class district is one which has a population of eight thousand or more, employs a superintendent who has had at least five years experience in public school work, and is administered by a board of seven trustees.

The second class district is defined as follows:[3]

> A second class district is one which has a population of one thousand or more and less than eight thousand; it employs a superintendent who has had at least three years experience in public school work; it is controlled by a board of five members.

The third class district is defined as follows:[4]

> A third class district is one with a population of less than one thousand and is controlled by a board of three members. A third class district employing more than one teacher may employ either a superintendent or principal or both. The one-room rural schools, of which there are slightly over nine hundred with an enrollment of approximately 9,300, are not included here.

The county high school is defined as follows:[5]

> A County High School is under the control of a County High School Board. The administrative officer for the board is the High School Principal. The County High School Board consists of seven members, of which the county superintendent is one and the remaining six are appointed by the county commissioners.

---

[2] Ibid., p. 5.

[3] Ibid., p. 8.

[4] Ibid., p. 17.

[5] Ibid., p. 27.

The questionnaire was mailed in the spring of 1953. Each teacher was requested to return, with the questionnaire, a major test, together with the pupil scores on the test. The pupils were to be designated only by letter or number.

Returns on the first distribution of the questionnaire were received from only thirty-eight teachers of the 100 to whom they were sent. The names of the elementary algebra instructors who did not reply to the first set of questionnaires were obtained from the Part A Forms in the State Department of Public Instruction in Helena through the courtesy of Mr. William I. King, State High School Supervisor. The elementary algebra instructors were not addressed by name when the first set of questionnaires was sent out. The Part A Forms revealed that ten third class schools apparently did not offer elementary algebra. None of these ten schools replied in the first survey. These ten schools were eliminated from the survey and ten other third class schools listed in the Part A files alternately to those eliminated were chosen. A second set of questionnaires was sent out to the sixty-two remaining schools. When only fifteen replied, follow-up letters were sent out and seventeen more replies were obtained.[6]

---

[6]The questionnaire, the letters mailed with the questionnaires in the two surveys, the follow-up letter, and the list of the schools in the survey may be found in the appendix.

The following table gives the number and percentage of questionnaire returns from each of the three classes of districts and the county high schools. It also indicates the number and percentage of teachers who sent tests, as well as the number and percentage of teachers who included test scores.

TABLE I

NUMBER AND PERCENTAGE OF RETURNS
ON THE
QUESTIONNAIRE, TESTS, AND TEST SCORES

	First Class*	Second Class*	Third Class*	County High School	All
Total number	5	67	106	17	195
Number in survey	5	35	55	5	100
Number of returns on questionnaire**	4	29	33	4	70
Percentage of returns on questionnaire	80.0	82.9	60.0	80.0	70.0
Number of teachers who returned tests	3	19	17	3	42
Percentage of returns on test	60.0	54.3	30.9	60.0	42.0
Number of teachers who returned pupil scores	0	8	11	2	21
Percentage of returns on pupil scores	0.0	22.9	20.0	40.0	21.0

*The county high schools are not included.

**Final returns

CHAPTER II

A REVIEW OF RELATED RESEARCH IN EXAMINATIONS AND MARKS OF TEACHERS

Various teachers tend to differ in their standards of marking. One teacher might judge a piece of work as being of a high quality where another teacher would judge the same work as being below par. W. S. Monroe reports such an investigation, made in 1913 by F. J. Kelly, of the marks given to the sixth-grade pupils in four ward schools in Hackensack, New Jersey, and the marks given to the same pupils when they went to a common departmental school for seventh-grade work.[1] Mr. Kelly stated this conclusion:

> "The work for which a teacher in school "C" (one of the ward schools) would give a mark of "G" (Good) in language, penmanship, or history, the teacher in school "D" (another ward school) would give less than a mark "F" (Fair)."

Mr. Kelly found that the pupils would receive a different kind of mark in the seventh grade common departmental school. This would indicate that there was a lack of consistency in the marking by ward school teachers.

Some teachers might display the tendency to give few high marks and many low marks. Other teachers might

---

[1]Walter Scott Monroe, Measuring the Results of Teaching. (Boston: Houghton Mifflin Company, 1918), pp.3-4.

tend to use many high marks and few lower marks. F. W. Johnson, then Principal of the University High School of the University of Chicago, made an investigation along this line.[2] Mr. Johnson bases his study on the fact that when accurate measurements are made of any ability of a large group of pupils, the resulting measures are distributed in the form of the normal curve. This curve is lowest at the "ends" and the highest toward the "middle", showing that more pupils tend to make "average" scores than high or low ones. Mr. Johnson noted in the University High School that more lower marks and few higher ones were given in English than in history. This same variation was found among two teachers in the same department. The teachers who give too high or too low marks conform less to the normal curve and are possibly less accurate in their marking.

An examination that is objective should permit little variation in answers, if any at all. A mathematics examination paper should be simple to grade objectively if mathematics, as is commonly conceived, is an exact subject. Professors Starch and Elliott carried on experiments along this line at the University of Wisconsin in 1912 and 1913. They studied the results of teacher-marking on two English papers, a history paper, and a geometry paper. They worked

---

[2]F. W. Johnson, "A Study of High School Grades", School Review, 19:13-24, January, 1911.

with two English papers selected from those written at the end of the first year of work in one of the large high schools of Wisconsin.[3] Plates were made of both these papers from which facsimiles were printed in as large numbers as necessary. The set of essay-type questions and a copy of each paper were sent to about 200 high schools. A request was made that the principal teacher of first-year English grade the two papers according to the practices and standards of the school. Replies came from approximately 150 teachers, ninety-one of whom were in high schools where the passing grade was seventy-five per cent. These ninety-one teachers graded one of the papers from sixty to ninety-seven per cent and the other from fifty to ninety-eight per cent.

A corresponding experiment was carried on in connection with a history examination.[4] Seventy history teachers in high schools marked a paper with seventy-five per cent as the passing grade. These teachers graded the history paper from forty-three to ninety-two per cent.

A similar procedure was followed in the grading of

---

[3]Daniel Starch and Edward C. Elliott, "The Reliability of Grading High School Work in English", School Review, 20:442-457, September, 1912.

[4]Daniel Starch and Edward C. Elliott, "The Reliability of Grading Work in History", School Review, 21: 676-681, December, 1913.

a geometry paper.[5] A copy of a reproduction of an actual examination paper in plane geometry was sent to each of the high schools included in the North Central Association of Colleges and Secondary Schools. The request was made that the paper be marked on a basis of 100 per cent being perfect. The teacher was asked to mark the paper by the method he was accustomed to using. The teachers from the 116 schools replying marked the geometry paper from twenty-eight to ninety-two per cent.

A teacher might judge a problem to be difficult and be more lenient in his marking policy as a result. Another teacher might consider the same problem to be relatively easy and therefore might not be as lenient in marking the problem. Mr. Comin submitted a list of twenty-three problems to twenty teachers who were asked to estimate the per cent of pupils who would solve each problem correctly if given ten minutes for each.[6] The teachers ranked the problems in order of difficulty according to their estimates. The same problems were given to the pupils in the fifth, sixth, seventh, and eighth grades in one school. The pupils' scores were used to rank the problems in order of difficulty, This rank of difficulty was compared with the

---

[5]Daniel Starch and Edward C. Elliott, "The Reliability of Grading Work in Mathematics", School Review, 21: 254-254-259, April, 1913.

[6]Robert Comin, "Teachers' Estimates of the Ability of Pupils", School and Society, 3:67, January 8, 1916.

teachers' rank of difficulty with the result that many discrepancies were found. A teacher may mark a problem harshly thinking the problem should be easy for the pupil when it actually is not.

A teacher's scheme for assigning grades may be so much a part of his professional character that much opposition is met by the principal or superintendent who attemps to secure a more uniform marking practice throughout the entire school or school system. Don C. Bliss made a report of the failure marks given during one semester in five of the large high schools of northern New Jersey. The tabulations in the study were made under the direction of Mr. M. R. Trabue by Mr. J. K. Walsh and other members of a class in educational administration at Teachers College.[7]

In two of the schools in the study only eight per cent of the marks indicated failure. In the third school twelve per cent of the marks were failure; in the fourth school, fourteen per cent were failing; and in the fifth school, twenty-seven per cent were failures. These schools were alike in size and ssumed to be alike in the intellectual abilities of the pupils. The five schools were in a similar environment. The teachers in these schools had met about the seme requirements for certification. The teachers in the school with the highest per-

---

[7]Marion Rex Trabue, Measuring Results in Education. (New York: American Book Company, 1924),pp. 43-49.

centage of failures assigned many failures in the third-year class as well as in the first two. These teachers also gave some failures in the fourth-year class. The other four schools followed the tendency of failing few third- and fourth-year students. There may have been teacher attitudes that "only the best pupils are able to pass the high school subjects". Pupils passing in one school might have failed in another.

The marks assigned by teachers of the same subject in the same high school were noted in this study to differ. The high school which had reported twenty-seven per cent failures showed the greatest amount of variation in marks within its own ranks as well as the greatest amount of deviation from the other high schools with which it was compared. If it is assumed that all the pupils under these various teachers did not differ too much in their ability and in what they could accomplish, the wide differences are probably due to the differences in the standards used by these teachers in making their judgments. A grade of A given by one teacher seldom giving A's might have more meaning than an A grade given by another teacher frequently giving A's.

B. D. Wood has shown the inadequacies of marks given in correcting college entrance examinations in algebra and geometry in June, 1921. About 400 algebra and 400 geometry papers chosen at random were each scored

twice independently by two different readers of the board. According to Mr. Wood, the scoring was such that agreement on failures would be somewhat as follows:

> "In gross numbers suppose that 10,000 candidates are tested with form A of a given geometry examination and that 30% of the 10,000 fail. The same 10,000 are tested with another equivalent geometry examination, form B, for example, which also failes 30% of the candidates. Accepting the results of one such examination as valid, which was done by the Board in 1921, another equivalent examination would pass 1,279 of the 3,000 failed by the first, and would fail 1,279 of the 7,000 passed on into college by the first. The fate of over 2,500 in 10,000 candidates would be reversed by an equivalent form of the same examination when the reliability is no higher than that of the College Entrance Examination Board Mathematics C Examination for June, 1921."[8]

Successive College Entrance Examination Board examinations were found to be unequal in difficulty. Mr. Wood made the following observation in this respect:

> "Large and unrecognized differences were found between the difficulties of the examinations which were thought to be equal. According to the reports of the Secretary of the College Entrance Examination Board, the percentage of failure in algebra in 1916 was 61.8, in 1917 was 36.7, in 1918 was 25.3, in 1919 was 61.3, in 1920 was 26.1, and in 1921 was 28.5."[9]

Mr. Wood also referred to a model history paper written by one of the expert readers assigned to score a certain group of history papers. The paper accidentally

---

[8]B. D. Wood, Measurements in Higher Education. (New York, Yonkers-on-Hudson: World Book Company, 1923), pp. 124-125.

[9]Ibid., p. 128 f.

came into the possession of another reader who proceeded to grade it. He assigned a below-passing grade to the paper. The other expert readers graded this paper from 40 to 90.

School marks were meant to be used only as a measure of achievement. Marks would be less meaningful if they were to deviate from that purpose. F. S. Camp held faculty meetings in his school in Stamford, Connecticut, to discuss the bases for giving school marks.[10] He found that teachers measured such attributes as improvement, ability, serious purpose, moral qualities, interest in work and accomplishments. One teacher in the mathematics department marked on "acquisition of knowledge". Marks had come to be used as a reward for the well-behaved student and as a disciplinary weapon to control the more troublesome student. Mr. Camp concluded that marks are almost meaningless if they measure many other qualities beside achievement.

W. H. Hughes found that teachers in one high school and junior college considered some thirty-four traits or factors when marking pupils.[11] Mr. Hughes also found that pupils were penalized because the teacher did not think

---

[10]F. S. Camp, "Some 'Marks': an Administrative Problem," School Review, 25:697-713, December, 1917.

[11]W. Hardin Hughes, "Coordinating Subject Marks and Achievement Marks," Nations Schools, 10:25-28, October, 1932.

them industrious or cooperative.[12]

The sex of the pupil might contribute to a difference in marking. A study by J. M. Lee and Dorris Lee revealed such a situation.[13] The pupils were grouped according to various achievement levels as determined by an achievement test and then separated by sex. The grade-point averages in both geometry and algebra were determined for each sex in each achievement level. The girls had an average of 0.656 grade-points better in algebra and an average of 0.240 grade-points better in geometry. In one case the girls were given a better mark for lower achievement. The whole situation was difficult to account for as most of the instructors were women.

Lee and W. H. Hughes collected data showing a difference in marking practices found in schools in the same city.[14] The first school gave forty-six per cent A's, as compared with twenty-two per cent A's in the second school, and thirteen per cent A's in the third school. The median achievement of the A pupils in the first school was one point lower than the median achievement of the D pupils in the third school. In the first

---

[12]W. Hardin Hughes, "Analyzing the Ingredients of Teachers' Marks," Nation's Schools, 6:21-25, December, 1930.

[13]J. Murray Lee, A Guide to Measurement in Secondary Schools. (New York: D. Appleton-Century Company, Inc., 1936), pp. 231-232.

[14]Ibid., pp. 232-235.

school the median achievements of the B and F pupils were higher than the median achievement of the A pupils in the same school.

Lee found variations in marking practices among six mathematics teachers in a certain junior high school.[15] One teacher assigned as low as seven per cent A's while another teacher assigned as high as twenty-five per cent A's. The assignment of B grades ranged from twelve to thirty-seven per cent, C grades from eighteen to fifty-four per cent, D grades from nine to twenty-two per cent, and E grades from zero to nine per cent. The differences in pupil ability might have accounted for some of these variations.

Mr. Lee also made an investigation of the amount of agreement among teachers in the marking of a test paper.[16] An arithmetic paper was given to a number of teachers to mark on the basis of 100 percent as perfect with each problem worth twenty per cent. Marks ranging from ten to sixty per cent were assigned by thirty-two experienced teachers in a tests and measurements class. Twelve of these teachers gave marks between thirty and thirty-nine per cent. The median score was thirty-six per cent.

---

[15] Ibid., pp. 236-237.

[16] Ibid., pp. 256-258.

Part of this survey will be devoted to a study of the marking practices of Montana elementary algebra teachers. The study should help to show the amount of agreement or disagreement in grading of algebra problems among Montana teachers.

Mr. Comin's study showed that the teachers' ideas of the difficulty of a problem differed from that of the pupils'.[17] The study of the writer will show the variation of the amount of difficulty of elementary algebra tests of different teachers. A teacher may give a test thinking it to be of moderate difficulty or even easy, when actually the test is quite difficult as far as the pupils are concerned. Algebra tests in the same category, e. g., six weeks by different teachers, probably differ in difficulty because teachers differ among themselves in their ideas of what constitute difficulty.

---

[17]Comin, loc. cit.

CHAPTER III

AN ANALYSIS OF INFORMATION COMPILED FROM THE QUESTIONNAIRE

A. GENERAL INFORMATION

A questionnaire was sent to Montana elementary algebra instructors to aid in the study of factors lying within the teacher that affect tests. The first section of the questionnaire pertains to general information about the schools from which questionnaire returns were received. The following table gives the enrollment of the secondary schools in the study and the break-down of enrollment according to districts and county high schools.

TABLE II

ENROLLMENT OF SECONDARY SCHOOLS ACCORDING TO DISTRICT BY FREQUENCY OF RESPONSES OBTAINED ON THE QUESTIONNAIRE

District	1000 or more	251-500	100-250	Less than 100	No Reply	Total
First Class	3	1	-	-	-	4
Second Class	-	3	19	4	3	29
Third Class	-	-	1	31	1	33
County High School	2	2	-	-	-	4
All	5	6	20	35	4	70

The schools in the questionnaire study were also broken down according to their location in the state of Montana. These geographical sections (southwest, northwest, northeast, and southeast) were arbitrarily formed by the intersection of the 47th parallel with the 110th degree of longitude. The following table gives the enrollment of the same secondary schools according to geographical sections.

TABLE III

ENROLLMENT OF SECONDARY SCHOOLS
ACCORDING TO GEOGRAPHICAL SECTION
BY FREQUENCY OF RESPONSES OBTAINED ON THE QUESTIONNAIRE

Geographical Section	1000 or more	251-500	100-250	Less than 100	No Reply	Total
Southwest	2	4	4	9	-	19
Northwest	2	2	8	8	-	20
Northeast	-	-	3	7	2	12
Southeast	1	-	5	11	2	19
All	5	6	20	35	4	70

The next table gives the number of elementary algebra sections taught by each teacher in the study and the number of sections taught by each teacher according to districts.

TABLE IV

NUMBER OF SECTIONS OF ELEMENTARY ALGEBRA TAUGHT ACCORDING TO DISTRICT BY FREQUENCY OF RESPONSES OBTAINED ON THE QUESTIONNAIRE

District	5	4	3	2	1	No Reply	Total
First class	1	2	-	1	-	-	4
Second class	-	-	6	9	14	-	29
Third class	-	-	1	2	29	1	33
County High School	-	-	2	2	-	-	4
All	1	2	9	14	43	1	70

The following table gives the number of students taught by each algebra teacher according to districts.

TABLE V

TOTAL NUMBERS OF STUDENTS TAUGHT ACCORDING TO DISTRICT BY FREQUENCY OF RESPONSES OBTAINED ON THE QUESTIONNAIRE*

District	Less than 5	5-10	11-15	16-20	21-30	31-50	More than 50	Total
First Class	-	-	-	-	-	-	4	4
Second Class	-	-	-	3	10	8	8	29
Third Class	3	7	13	2	6	2	-	33
County High School	-	-	-	-	-	-	4	4
All	3	7	13	5	16	10	16	70

*In some cases the total number of students was contained in more than one section.

The number of years of teaching experience of each algebra teacher is given in the following table indicated by frequency of responses.

TABLE VI

TEACHING EXPERIENCE IN YEARS BY DISTRICT
BY FREQUENCY OF RESPONSES OBTAINED ON THE QUESTIONNAIRE

District	0	1	2	3	4	5	6-10	11-15	16-20	21-30	More than 30	Total
First class	-	-	-	1	1	-	-	-	-	1	1	4
Second class	2	2	1	5	3	1	2	5	2	4	2	29
Third class	2	1	7	4	2	1	6	2	2	3	2	32*
County High School	-	-	-	-	-	1	-	-	-	2	1	4
All	4	3	8	10	6	3	8	7	4	10	6	69

*One third class teacher did not reply to this item.

Each teacher was also asked to give the number of years he had been employed in his present position. The following table gives this information by districts with the frequency of responses indicated.

TABLE VII

YEARS OF TEACHING EXPERIENCE IN PRESENT POSITION BY DISTRICT
BY FREQUENCY OF RESPONSES OBTAINED ON THE QUESTIONNAIRE

District	0	1	2	3	4	5	6-10	11-15	16-20	More than 20	Total
First class	-	1	1	-	-	-	1	-	-	-	3*
Second class	2	8	3	3	-	3	6	3	-	1	29
Third class	5	16	4	2	3	2	-	1	-	-	33
County High School	-	-	-	-	-	1	1	1	-	1	4
All	7	25	8	5	3	6	8	5	-	2	69

*One first class teacher did not reply to this item

The following table gives a list of the elementary algebra textbooks used by the teachers in this study.

TABLE VIII

ELEMENTARY ALGEBRA TEXTBOOKS IN USE

Author	Title	Publisher	Copyright Date	Frequency of Use
Welchons and Krickenberger	Algebra, Book One	Ginn and Company	1949 1953	13 1
Edgerton and Carpenter	Elementary Algebra	Allyn and Bacon	1942 1943 1945 1947 1949 1951	1 1 1 2 7 2
Hawkes, Luby and Touton	First Year Algebra	Ginn and Company	1934 1946 1949 1951	1 3 1 3
Wells and Hart	Progressive First Algebra	D. C. Heath and Company	1934 1943	3 3
Walter W. Hart	A First Course in Algebra	D. C. Heath and Company	1947 1951	3 2
Freilich, Berman, and Johnson	Algebra for Problem Solving	Houghton, Mifflin Company	1952	3
Stein	Algebra in Easy Steps	Van Nostrand	1949	3
Nyberg	Fundamentals of Algebra	American Book Company	1940 1944	1 2
Walter W. Hart	Essentials of Algebra	D. C. Heath and Company	1941	2
Hawkes, Luby and Touton	Elementary Algebra	Ginn and Company	1942	2
D. W. Snader	Elementary Algebra	John C. Winston Company	1949	2

TABLE VIII (continued)

ELEMENTARY ALGEBRA TEXTBOOKS IN USE

Author	Title	Publisher	Copyright Date	Frequendy of Use
William Betz	Algebra for Today	Ginn and Company	1937 1941	1 1
William Betz	Everyday Algebra	Ginn and Company	1950	1
William Betz	Everyday Junior Mathematics	Ginn and Company	1946	1
Ernest R. Breslich	Algebra First Course	Laidlaw Bros.	1943	1
Foster E. Grossnickle	General Mathematics	John C. Winston Company	1949	1
Walter W. Hart	Elementary Algebra	Not given	1940	1
Walter W. Hart	First Year Algebra	Ginn and Company	1952	1
Lennes	A First Course in Algebra	Macmillan	1950	1
Milne	High School Algebra	American Book Company	1920	1
Stone-Mallory	A First Course in Algebra	Sanborn	1936	1
Upton	Algebra	American Book Company	1936	1
			Total	74

A total of seventy-four responses was received on this item.

Two teachers used three different textbooks and two teachers used two different textbooks. Two teachers did not reply to this item.

Minnick gives a summary of the qualities to be found in a suitable textbook.[1] He states that a textbook should have an attractive design related to the content of the book. The binding of the textbook should be durable. The pages in the book should present a neat appearance. Ample margins and proper line spacing are desirable. The type should be easy to read with the avoidance of too much heavy type. The textbook material should be orderly to enable the pupils to gain a more orderly picture of the material. The table of contents should give a general view of the entire book and the index should make it possible to find quickly any desired details. The applications of the subject matter should arouse the pupil's interest, make the pupil conscious of the usefulness of the mathematical facts learned, and enable the pupil to use his mathematical knowledge and skills when the occasion arises. Drill materials should be selected on the basis of useful application. The arrangement of subject matter should show the usefulness of the material. Adequate drill material can be found in the supplementary exercises of a good textbook. Lastly, a properly arranged textbook can

---

[1]G. H. Minnick, Teaching Mathematics in the Secondary Schools - Principles and Methods. (New York: Prentice-Hall, Inc., 1939), pp. 276-287.

show the relation of the subject to the progress of civilization.

The following table gives a rating on the basis of excellent, good, fair, and poor of a sampling of textbooks typical of those used by Montana elementary algebra teachers. These books were rated by the writer according to the qualities set forth by Minnick.

TABLE IX

RATINGS OF ELEMENTARY ALGEBRA TEXTBOOKS USED BY MONTANA ELEMENTARY ALGEBRA TEACHERS

Qualities designated by Minnick	Welchons and Krickenberger Algebra, Book One	Edgerton and Carpenter-Elementary Algebra Copyright, 1949	Walter W. Hart A First Course in Algebra Copyright, 1947	Freilich, Berman and Johnson Algebra for Problem Solving Copyright, 1952	Ernest R. Breslich Algebra First Course Copyright, 1943
Design	Good	Excellent	Good	Excellent	Fair
Binding	Excellent	Good	Excellent	Excellent	Fair
Margins	Good	Good	Good	Excellent	Fair
Line spacing	Excellent	Excellent	Excellent	Excellent	Excellent
Type	Excellent	Excellent	Excellent	Excellent	Excellent
Orderliness and subject matter arrangement	Good	Good	Good	Good	Fair
Table of contents	Excellent	Fair	Fair	Good	Fair
Index	Excellent	Fair	Good	Excellent	Excellent
Subject matter applications	Good	Good	Good	Excellent	Fair
Drill material and supplementary exercises	Excellent	Good	Excellent	Excellent	Good
Cultural use	Excellent	Good	Excellent	Excellent	Fair

Each teacher in the survey was requested to state whether or not elementary algebra is a graduation requirement in the high school in which he taught. The following table gives the results on this item according to districts and to geographical sections.

TABLE X

IS ELEMENTARY ALGEBRA A GRADUATION REQUIREMENT?

District	Yes	No	Total	Geographical Section	Yes	No	Total
First class	-	4	4	Southwest	7	12	19
Second class	10	19	29	Northwest	6	14	20
Third class	25	8	33	Northeast	8	4	12
County High School	-	4	4	Southeast	14	5	19
All	35	35	70	All	35	35	70

The smaller high schools tend to require elementary algebra for graduation, while the larger high schools and the county high schools, two of which in this study have large enrollments, tend to drop this requirement. Section-wise, the eastern part of the state tends to retain this requirement, while the western part tends to permit graduation without elementary algebra. All the schools taken together are evenly split on this item.

Each teacher was asked to state whether or not any standardized prognostic elementary algebra tests are given

to the students before they take the course. The following table gives the results on this item according to districts and to geographical sections.

TABLE XI

ARE STANDARDIZED PROGNOSTIC ELEMENTARY ALGEBRA TESTS GIVEN?

District	Yes	No	Total	Geographical Section	Yes	No	Total
First class	1	3	4	Southwest	5	14	19
Second class	9	19	28*	Northwest	7	13	20
Third class	2	31	33	Northeast	1	10	11**
County High School	2	2	4	Southeast	1	18	19
All	14	55	69	All	14	55	69

*One second class teacher did not reply.

**One northeast teacher did not reply.

The movement toward this kind of test tends to be stronger in the larger high schools and in the western part of the state. The next table gives the names of the prognostic elementary algebra tests in twelve out of fourteen cases where this kind of test was administered. Two teachers did not reply to the name of the test they gave.

TABLE XII

PROGNOSTIC ELEMENTARY ALGEBRA TESTS GIVEN WITH FREQUENCY OF QUESTIONNAIRE RESPONSES ACCORDING TO DISTRICT

	First Class	Second Class	Third Class	County High School	All
California Algebra Aptitude Test	-	3	-	-	3
Iowa Algebra Aptitude Test	-	1	1	1	3
Lee Test of Algebraic Ability	-	1	1	1	3
Orleans Algebra Prognosis Test	-	2	-	-	2
Test from Social Utility Arithmetic, Book V	-	1	-	-	1
Total	-	8	2	2	12

Two types of procedure have been tried as a basis for prognosis.[2] One of these is the learning technique, in which the aptitude of the student is measured in terms of the speed and accuracy with which he is able to acquire skills and information and respond to objective tests over the newly learned material; the other is the inventory technique, in which the student reveals his aptitude in terms of reactions to specific exercises sampling into underlying skills upon which success in the subject depends. The Orleans Algebra Prognosis Test embodies the learning technique, while the Iowa Algebra Aptitude Test (Revised), the Lee Test of Algebraic Ability, and the California Algebra Aptitude Test are of the inventory type. The Iowa Algebra Aptitude Test (Revised) selects four basic skills dealing with arithmetic computations, computations of abstract concepts, manipulation of numerical series, and problems of dependence and variation. The test by Orleans is divided into a test on arithmetic and into twelve other parts.[3] They are: substitution in monomials, use of exponents, meaning of exponents, substitution in monomials with exponents, substitution in binomials with exponents,

---

[2]Harry A. Greene, Albert N. Jorgensen, and J. Raymond Garberich, Measurement and Evaluation in the Secondary School. (New York: Longmans, Green and Company, 1943), pp. 402-403.

[3]A. M. Jordan, Measurement in Education. (New York: McGraw-Hill Book Company, Inc., 1953), pl 241

like and unlike terms, representation of relations, representation of expressions, positive and negative numbers, problems, addition of like terms, and a summary test. This test is now longer and the validity coefficient is about 0.80.

In some cases standardized elementary algebra tests other than prognostic were given to the students. The results on this questionnaire item are given in the following table according to districts and to geographical sections.

TABLE XIII

ARE STANDARIZED ELEMENTARY ALGEBRA TESTS OTHER THAN PROGNOSTIC GIVEN?

District	Yes	No	Total	Geographical Section	Yes	No	Total
First class	2	2	4	Southwest	8	11	19**
Second class	5	22	27*	Northwest	5	14	19**
Third class	7	26	33	Northeast	3	8	11**
County High School	3	1	4	Southeast	1	18	19
All	17	51	68	All	17	51	68

*Two second class teachers did not reply.

**One northwest and one northeast teacher did not reply.

The tendency toward the use of these tests is stronger in the western part of the state. The following table gives names of these algebra tests. The number of responses totals

eighteen as one school gave both the Seattle Algebra Test and the Lankton and another school gave both the Every Pupil Scholarship Test and the High School Content. However, one teacher did not give the name of the test given.

TABLE XIV

STANDARDIZED ELEMENTARY ALGEBRA TESTS OTHER THAN PROGNOSTIC BY FREQUENCY OF QUESTIONNAIRE RESPONSES BY DISTRICT

	First Class	Second Class	Third Class	County High School	All
Every Pupil Scholarship Test, Bureau of Educational Measurements, Kansas State Teachers College (given in January and April	-	1	5	1	7
Seattle Algebra Test (Given at the end of the first semester)	-	2	-	1	3
Cooperative Algebra Test Long, Siceloff, and Chesire, Educational Testing Service (Given in late spring)	1	1	-	-	2
Larson-Greene Unit Tests in First Year Algebra (Given at completion of units of work)	1	-	-	1	2
High School Content (Given in May)	-	-	1	-	1
Iowa Algebra Achievement Test (Given near end of school year)	-	1	-	-	1
Lankton Test book with textbook (Given at end of school year)	-	-	1	-	1
Total	2	5	7	4	18

B. THE PREPARATION OF TEACHER-MADE MAJOR TESTS

The major tests are the examinations given at the end of six or nine weeks' period, the semester, and the end of the school year. The teachers in this survey were found to vary considerably in the number of these tests they gave in a school year. The following table gives a general picture of the number of these tests given.

TABLE XV

THE NUMBER OF MAJOR ELEMENTARY ALGEBRA TESTS IN A SCHOOL YEAR BY FREQUENCY OF RESPONSES OBTAINED ON THE QUESTIONNAIRE

District	2	3-5	6	7	8	9-12	13-20	More than 20	Total
First class	-	-	1	-	-	1	-	1	4
Second class	5	2	7	-	10	2	3	-	29
Third class	2	1	12	1	9	1	5	2	33
County High School	-	-	-	-	2	1	1	-	4
All	7	3	20	1	21	5	10	3	70

The tendency is for the teacher to give six or eight major tests in a school year. The former plan would probably include two six weeks' tests and a final per semester and the latter, three six weeks' tests and a final per semester. The teachers in the larger high schools tend to give more tests a year than those in the smaller high schools.

Most teachers informed their students that a major test was coming. The next table gives the frequency of responses on this questionnaire item.

TABLE XVI

ARE STUDENTS INFORMED OF COMING MAJOR TESTS?

District	Yes	No	Sometimes	Total
First class	3	-	-	3*
Second class	28	-	1	29
Third class	30	2	1	33
County High School	4	-	-	4
All	65	2	2	69

*One first class teacher did not reply.

Informing students of a final examination to be given has a favorable effect on their learning. Lee cites an experiment on the college level showing that those who expected the final examination did markedly better on it than those who did not expect it.[4] Such a case might be true for any examination.

Many teachers tend to select their major test items from the basal textbooks used. The basal textbook is the principal book used by the student. The following table gives the frequency of responses on this item.

---

[4]J. Murray Lee, A Guide to Measurement in Secondary Schools. (New York: D. Appleton-Century Company, Inc., 1936), p. 327.

TABLE XVII

ARE MAJOR TEST ITEMS SELECTED FROM BASAL TEXTBOOKS?

District	Yes	No	Sometimes	Total
First class	1	3	-	4
Second class	17	7	5	29
Third class	29	2	2	33
County High School	2	2	-	4
All	49	14	7	70

There is a tendency for the high school teachers from the smaller schools to select their major test items from the basal textbooks used. The teachers from the two County High Schools with first class enrollments did not select their major test items from the basal textbooks used. The other two teachers are from County High Schools with second class enrollments.

Teachers were found to vary considerably in the time they took to prepare major tests. The following table gives the approximate time each teacher reported.

TABLE XVIII

TIME BY MINUTES SPENT IN PREPARATION OF MAJOR TESTS BY FREQUENCY OF QUESTIONNAIRE RESPONSES

District	30	45	60	75	90	120	More than 120	Total
First class	-	2	1	-	-	-	-	3*
Second class	1	3	7	1	4	6	7	29
Third class	5	6	8	-	6	2	6	33**
County High School	1	-	1	-	-	1	-	3*
All	7	11	17	1	10	9	13	68

*One first class and one County High School teacher did not reply.

**One third class teacher spent twelve hours in preparing a major test.

C. THE PREPARATION OF TEACHER-MADE MINOR TESTS

The minor tests are the quizzes given by the teacher daily or less often. Each teacher was asked to state how many of these quizzes he gave in a typical school week. Variation among teachers was not great on this questionnaire item. The following table gives the results by frequency of questionnaire responses.

TABLE XIX

THE NUMBER OF ELEMENTARY ALGEBRA QUIZZES GIVEN IN A TYPICAL SCHOOL WEEK

District	Less than 1	1-2	3-4	5	Varies	Total
First class	1	2	-	-	-	3*
Second class	3	20	4	1	1	29
Third class	3	23	2	1	1	30**
County High School	1	2	1	-	-	4
All	8	47	7	2	2	66

*One first class teacher did not reply.

**Two third class teachers gave no minor tests and one third class teacher did not reply.

Lee goes on to say that the use of a number of short tests throughout a course seemed to give better results than when such frequent tests were not used. The score on any one of these tests would be neither valid nor reliable. The combination of all the scores from the tests would give a valid and reliable measure if each test were valid and reliable.[5] The teacher who gives five quizzes a week could possibly have a better measurement than the teacher giving one or two quizzes a week. Lee also found that students preferred to have a number of short tests throughout the

[5] Loc. cit.

course. These students were concerned as to what type of progress they were making in their work.

Considerably fewer teachers informed their pupils of coming minor tests than was the case with major tests. Teachers have felt that students would prepare their lessons more diligently if they knew that a test could occur. Most pupils would be expecting a test which, according to Lee, would give the test greater instructional value. The more frequently these tests are given, the more the students are going to be prone to expect one. This is another advantage of frequent, short tests. The following table gives the results on this questionnaire item by frequency of responses.

TABLE XX

ARE STUDENTS INFORMED OF COMING MINOR TESTS?

District	Yes	No	Sometimes	Total
First class	2	1	-	3*
Second class	13	12	4	29
Third class	14	14	2	30*
County High School	-	4	-	4
All	29	31	6	66

*One first class and one third class teacher did not reply.

The tendency for teachers to select their minor test items from basal textbooks used is greater than was the case with major tests. The next table gives the results on this questionnaire item.

TABLE XXI

ARE MINOR TEST ITEMS SELECTED FROM BASAL TEXTBOOKS?

District	Yes	No	Sometimes	Total
First class	1	3	-	4
Second class	22	4	3	29
Third class	30	1	-	31
County High School	3	1	-	4
All	56	9	3	68

As with major tests, the smaller high school teachers tend more to select their minor test items from basal textbooks used.

Teachers did not vary as markedly in the time they took to prepare minor tests as with major tests. The approximate time each teacher took to prepare minor tests is indicated in the following table by frequency of responses.

TABLE XXII

TIME BY MINUTES SPENT IN PREPARATION OF MINOR TESTS BY FREQUENCY OF RESPONSES OBTAINED ON THE QUESTIONNAIRE

District	Less than 15	15	30	45	60	No Reply	Total
First class	-	1	2	-	-	1	4
Second class	6	9	7	2	1	4	29
Third class	10	8	6	3	2	2	31
County High School	3	1	-	-	-	-	4
All	19	19	15	5	3	7	68

D. ADMINISTRATION OF TEACHER-MADE MAJOR AND MINOR TESTS

After a teacher has prepared a test, he is ready to present it to his students. A test well-prepared can have little meaning as a measuring instrument if the procedures of administration are poor. In the first place the teacher should allow the students a fair amount of time to finish the test unless the examination is intended to measure speed of doing problems. The time allowed on the test itself should not be such that the students have too much idle time on their hands when they finish. The problem of major test length and time allowed are dealt with in the next chapter. Teachers did vary somewhat in the approximate time they allowed students for minor tests. The

following table gives the results on this questionnaire item by frequency of responses.

TABLE XXIII

TIME IN MINUTES ALLOWED STUDENTS FOR MINOR TESTS BY FREQUENCY OF QUESTIONNAIRE RESPONSES

District	Less than 15	15	30	45	60	No Set Time	Total
First class	-	2	2	-	-	-	4
Second class	5	9	8	3	1	3	29
Third class	6	10	9	4	1	-	30*
County High School	-	2	1	-	-	-	4
All	11	23	20	7	2	4	67

*One third class teacher did not reply. (The other two gave no minor tests.)

Twenty-six teachers in this survey had more than one section of elementary algebra. The majority of these teachers gave the same major and minor tests to each section of elementary algebra they taught. Under such a system the students of one section could pass on valuable information concerning the test to students in the following section. This inclination could be discouraged by curving the grades among all sections. Thus, a good performance by one section will push the passing grade higher. This would work a hardship on the first section without the

advantage of advance information about the test. However, some students for the sake of maintaining good standing with their friends, which is very important with adolescents, will give away test information even though they may suffer by it. A better solution to the problem is to give different major and minor tests to each section. The students could still tell what the general idea of the test was or what the important points consisted of. However, each problem in the different tests would not be exactly the same. The following tables give the results on this questionnaire item.

TABLE XXIV

IS THE SAME MAJOR TEST GIVEN TO EACH SECTION?

District	Yes	No	Total
First class	4	-	4
Second class	13	2	15
Third class	1	1	2*
County High School	3	1	4
All	21	4	25

*Three third class teachers had more than one section. However, one of these teachers had all the sections within one class at the same time.

TABLE XXV

IS THE SAME MINOR TEST GIVEN TO EACH SECTION?

District	Yes	No	Sometimes	Total
First class	4	-	-	4
Second class	10	4	1	15
Third class	-	2	-	2
County High School	1	3	-	4
All	15	9	1	25

There was a greater tendency to give different minor tests for each section. A minor test was generally shorter and an alternate form would be easier to prepare. The teachers in the larger high schools tend to give the same test to all sections. A teacher with five sections of algebra would as a rule find little time to prepare five separate tests.

Most teachers did not prepare alternate major or minor tests for use within a class or section. Under such a plan, each alternately seated student would be taking the same test. Thus, a student would find it difficult to look at his neighbor's paper from two seats away. The following tables give the results on this questionnaire item.

TABLE XXVI

ARE ALTERNATE MAJOR TESTS PREPARED?

District	Yes	No	Sometimes	Total
First class	-	4	-	4
Second class	3	24	2	29
Third class	2	29	-	31*
County High School	1	3	-	4
All	6	60	2	68

*Two third class teachers did not reply.

TABLE XXVII

ARE ALTERNATE MINOR TESTS PREPARED?

District	Yes	No	Sometimes	Total
First class	-	4	-	4
Second class	2	25	2	29
Third class	2	27	-	29*
County High School	1	3	-	4
All	5	59	2	66

*Two third class teachers did not reply. (The other two gave no minor tests.)

The most common form in which major tests are presented was by mimeographed sheets. Some teachers used the mimeographed sheets only for major tests, while others used mimeographed sheets on one occasion and the blackboard on another. Three or four different methods may have been used by one teacher for presenting major tests. The following table gives the results on this questionnaire item by frequency of responses by each teacher. The number of responses exceed seventy as s ome teachers presented their major tests in more than one form.

TABLE XXVIII

THE FORM OF PRESENTATION OF MAJOR TESTS
BY FREQUENCY OF QUESTIONNAIRE RESPONSES

District	Blackboard	Mimeographed Sheets	Dictation	Other Methods*	Total
First class	3	1	-	1	5
Second class	10	24	-	2	36
Third class	11	29	2	1	43
County High School	2	3	-	1	6
All	26	57	2	5	90

*Other methods, with frequency of responses in parentheses, are standardized tests (2), test booklets (1), hectograph (1), and not named (1).

Teachers in the larger high schools tend to use the blackboard more and those in the smaller schools, the mimeographed sheets. The greater number of students in the algebra classes in a first class school might make the blackboard method easier over producing a large number of mimeographed sheets. The student can work better with a mimeographed sheet at hand than by having to look up to the blackboard so often. However, the mimeographed sheet must be easy to read if it is to be an effective form for presenting a major test. The method of dictation may lead to misunderstanding on the part of the students. Those with hearing difficulties or those in the back of the room may be penalized. The method of dictation is inferior and probably should not be used.

The most common form, however, for the presentation of minor tests was the blackboard. The following table gives the results on this item by frequency of responses for the districts. The number of responses exceed sixty-eight, the number of teachers giving minor tests, as some teachers presented their minor tests in more than one form.

TABLE XXIX

THE FORM OF PRESENTATION OF MINOR TESTS
BY FREQUENCY OF QUESTIONNAIRE RESPONSES

District	Black-board	Mimeographed Sheets	Dicta-tion	Other Methods*	Total
First class	4	-	-	1	5
Second class	19	12	3	3	37
Third class	20	16	3	3	42
County High School	4	1	1	-	6
All	47	29	7	7	90

*Other methods, with frequency of responses in parentheses, are test booklets (4), hectograph (1), and not named (2).

Again the tendency for the teacher to use the blackboard is stronger in the larger high schools. The method of dictation is used more with minor tests than with major.

Many teachers followed the practice of assigning alternate seats during a major test. However, only a minority of teachers followed this procedure during a minor test. In some cases alternate seats could not be assigned because room did not permit. The following tables give the results on this questionnaire item.

TABLE XXX

ARE ALTERNATE SEATS ASSIGNED DURING A MAJOR TEST?

District	Yes	No	Room Does Not Permit	Sometimes	Total
First class	-	2	2	-	4
Second class	18	8	2	1	29
Third class	25	7	1	-	33
County High School	2	-	1	1	4
All	45	17	6	2	70

TABLE XXXI

ARE ALTERNATE SEATS ASSIGNED DURING A MINOR TEST?

District	Yes	No	Room Does Not Permit	No Reply	Total
First class	-	2	2	-	4
Second class	7	19	3	-	29
Third class	13	15	-	3	31
County High School	-	2	1	1	4
All	20	38	6	4	68

The larger the high school, the less likely the teacher was to assign alternate seats. This is probably because of lack of room, and in some cases the teacher replied that such is the situation. The majority of teachers did not solve their problems of independent work by alternate tests or by different tests for each section. The favorite plan was apparently the assignment of alternate seats when possible. The assignment of alternate seats, which was probably thought of as easier than the preparation of alternate tests and more practical, was probably just as effective. A student would need to make obvious gestures in trying to look at a paper from two seats away.

Most teachers did not feel that independent work on the part of their students during a test presented a problem. Apparently they felt that they had found some means to handle the situation. The following table gives the results on this item.

TABLE XXXII

DOES INDEPENDENT WORK PRESENT A PROBLEM DURING A TEST?

District	Yes	No	Sometimes	Total
First class	2	2	-	4
Second class	4	24	1	29
Third class	2	29	1	32*
County High School	-	4	-	4
All	8	59	2	69

*One third class teacher did not reply.

The teachers in the larger high schools might have more of a problem of independent work than those in the smaller high schools. The exception is with the County High Schools, two of the above which have first class enrollments. Independent work on the part of the students is probably going to be more of a problem under crowded conditions.

These teachers expressed varied opinions for the handling of the problem of independent work. These opinions were cataloged into three categories or approaches, the lenient, the moderate, and the strict. A teacher taking a lenient approach to this problem appeals to the honesty of the student or ignores the problem altogether. Those taking a moderate approach employ various measures to handle the problem such as alternate seats or alternate tests. The student is not completely trusted, but he

apparently is not threatened with firm measures to be taken if he is caught cheating. A strict approach to the problem indicates that firm measures or definite punishment will follow cheating, such as the failure of the student for that particular test. The following table gives the results as to these approaches by frequency of responses.

TABLE XXXIII

APPROACHES TO THE HANDLING OF THE PROBLEM OF INDEPENDENT WORK BY FREQUENCY OF QUESTIONNAIRE RESPONSES

District	Lenient Approach	Moderate Approach	Strict Approach	No Reply	Total
First class	1	3	-	-	4
Second class	4	10	8	7	29
Third class	2	13	2	16	33
County High School	-	2	1	1	4
All	7	28	11	24	70

The following teachers stated varied kinds of approaches to the problem:

One teacher took this kind of a lenient approach;

> "By not worrying too much about it. Some poor student might actually learn a little by copying."

Another teacher made an appeal to student honesty;

> "I usually appeal to their honor. I emphasize understanding, not grades."

One teacher took the following moderate approach;

> "First have confidence and respect of students. Then spread them way out and give special attention to security risks."

Another teacher simply said:

> "Alternating seats, alternate tests."

One teacher would fail the cheater;

> "Have the authority to pick up paper of cheating student and give said student an F for that test."

Again, another teacher simply said;

> "Stiff punishment to offenders."

Many teachers placed little to no emphasis on memorization of material in the testing program. They would rather emphasize understanding of material. However, some teachers felt that their students should know the quadratic formula by rote. The following table gives the results on this questionnaire item.

TABLE XXXIV

DEGREES OF EMPHASIS ON MEMORIZATION OF MATERIAL IN THE TESTING PROGRAM

District	Much	Some	Little	None	No Reply	Total
First class	-	-	1	2	1	4
Second class	1	10	14	3	1	29
Third class	2	6	14	9	2	33
County High School	-	2	1	1	-	4
All	3	18	30	15	4	70

Emphasis on memorization by second class teachers is slightly stronger than the emphasis by third class teachers.

E. THE EVALUATION PROCEDURES FOR MAJOR AND MINOR TESTS

The vast majority of teachers was against allowing students to correct their own major test papers. Many teachers did not allow the students to correct their own minor test papers. Again, nearly all the teachers did not allow the students to correct major test papers not their own. However, a fair share of teachers did allow the students to correct minor test papers not their own. The following tables give the results on these questionnaire items.

TABLE XXXV

ARE STUDENTS ALLOWED TO CORRECT THEIR OWN MAJOR TEST PAPERS?

District	Yes	No	Sometimes	Total
First class	-	4	-	4
Second class	-	28	1	29
Third class	1	32	-	33
County High School	-	4	-	4
All	1	68	1	70

TABLE XXXVI

ARE STUDENTS ALLOWED TO CORRECT THEIR OWN MINOR TEST PAPERS?

District	Yes	No	Sometimes	Total
First class	-	2	2	4
Second class	5	21	3	29
Third class	6	25	-	31
County High School	1	3	-	4
All	12	51	5	68

TABLE XXXVII

ARE STUDENTS ALLOWED TO CORRECT MAJOR TEST PAPERS NOT THEIR OWN?

District	Yes	No	Sometimes	Total
First class	-	4	-	4
Second class	-	29	-	29
Third class	1	30	2	33
County High School	-	4	-	4
All	1	67	2	70

TABLE XXXVIII

ARE STUDENTS ALLOWED TO CORRECT MINOR TEST PAPERS NOT THEIR OWN?

District	Yes	No	Sometimes	Total
First class	2	2	-	4
Second class	9	15	5	29
Third class	11	16	4	31
County High School	2	2	-	4
All	24	35	9	68

Lee finds that a test has a great instructional value to the pupil when he is allowed to correct his own paper.[6] Lee cites a study by Curtis and Woods, who found that students did better when tested the next day on material they themselves had corrected. These students also did better when tested at the end of six weeks. Their final gain was greater than the immediate gain by discussion of errors in class on teacher-corrected-papers. The study by Curtis and Woods also showed that pupils were able to correct their papers with a sufficiently high degree of accuracy to prove satisfactory.

Perhaps a great number of Montana teachers in this survey frowned on student correction of papers because they felt it was poor judgment to place such an opportunity for cheating in front of the pupils. It is a case of the measurement functions against the instructional functions of the same test. Lee suggests a solution to this problem by the double answer column or extra answer sheet. The student duplicates his answers on the extra column or extra sheet. The student hands in the torn-off extra column or extra answer sheet to the teacher, who then has a record of the answers. Lee has found the use of the extra answer sheet to be the most satisfactory plan. The teacher can hold a discussion period as the test is being corrected. This

---

[6] Ibid., pp. 328-331

plan can save the teacher much time in the correcting of papers. The teacher might choose better items for a test that is to be used as a learning device. However, such a plan requires more class time to give and correct the test, and is applicable only when each item has just one correct answer.

At the end of each report period the student is given a mark. A major test, such as the six weeks' test, is employed to help to derive that mark. The teachers in this survey differed somewhat in the weights they gave this major test. Many would stipulate that a six weeks' test counts twenty-five or thirty-three and one-third per cent. The following tables give the results on this questionnaire item for the major test by districts and geographical sections.

TABLE X XXIX

THE PERCENTAGE EACH MA JOR TEST COUNTS
TOWARD A STUDENT'S FINAL MA RK FOR A REPORT PERIOD
BY FREQUENCY OF RESPONSES OBTAI NED ON THE QUESTIONNAIRE

District	10%	16 2/3%	20%	25%	30%	33 1/3%	40%	50%	70%	No Reply	Total
First class	-	1	-	1	-	-	-	-	-	2	4
Second class	-	-	2	10	2	11	2	-	1	1	29
Third class	2	1	1	7	2	10	3	3	-	4	33
County High School	-	-	-	-	-	1	-	1	-	2	4
All	2	2	3	18	4	22	5	4	1	9	70

TABLE XL

THE PERCENTAGE EACH MAJOR TEST COUNTS
TOWARD A STUDENT'S FINAL MA RK FOR A REPORT PERIOD
BY FREQUENCY OF RESPONSES OBTAINED ON THE QUESTIONNAIRE

Geographical Section	10%	16 2/3%	20%	25%	30%	33 1/3%	40%	50%	70%	No Reply	Total
Southwest	2	1	1	4	2	4	1	-	1	3	19
Northwest	-	-	2	3	1	7	1	3	-	3	20
Northeast	-	1	-	2	-	5	1	1	-	2	12
Southeast	-	-	-	9	1	6	2	-	-	1	19
All	2	2	3	18	4	22	5	4	1	9	70

The following table gives the average percentage a major test counts by districts and geographical sections.

TABLE XLI

THE AVERAGE PERCENTAGE EACH MAJOR TEST COUNTS TOWARD A STUDENT'S FINAL MARK FOR A REPORT PERIOD

District	Average Percentage*	Geographical Section	Average Percentage*
First class	21.00	Southwest	29.00
Second class	31.00	Northwest	33.00
Third class	31.00	Northeast	32.00
County High School	42.00	Southeast	30.00
All	31.00	All	31.00

*Approximate figures

The teacher may have given a certain number of minor tests by the end of a report period. The minor tests were also used to help to derive a student's mark for that report period. The mark on one minor test or daily quiz is not so significant. However, all these marks taken together can produce a meaningful result. Teachers varied more in the weights they gave to minor tests than was the case for a major test. Again, the thirty-three and one-third per cent mark appears to be a popular weight. The following tables give the results on this questionnaire item by districts and geographical sections.

TAB LE XLII

THE PERCENTAGE ALL MINOR TESTS COUNT
TOWARD A STUDENT'S FIN AL MARK FOR A REPORT PERIOD
BY FREQUENCY OF RESPONSES OBTAINED ON THE QUESTIONNAIRE

District	10% or less	16 2/3%	20%	25%	27½%	30%	33 1/3%	35%	37½%	40%	44 4/9%	50%	60%	66 2/3%	75%	No Reply	Total
First class	1	-	1	-	-	-	-	-	-	-	-	1	-	-	-	1	4
Second class	2	1	1	4	-	1	7	1	1	1	1	2	1	1	1	4	29
Third class	4	1	4	1	1	2	5	1	-	1	-	3	-	-	1	9	33
County High School	1	-	-	-	-	-	-	-	-	-	-	1	-	1	-	1	4
All	8	2	6	5	1	3	12	2	1	2	1	7	1	2	2	15	70

TAB LE XLIII

THE PERCENTAGE ALL MINOR TESTS COUNT
TOWARD A STUDENT'S FIN AL MARK FOR A REPORT PERIOD
BY FREQUENCY OF RESPONSES OBTAINED ON THE QUESTIONNAIRE

Geographical Section	10% or less	16 2/3%	20%	25%	27½%	30%	33 1/3%	35%	37½%	40%	44 4/9%	50%	60%	66 2/3%	75%	No Reply	Total
Southwest	3	-	3	1	-	2	2	1	-	-	-	2	1	1	-	3	19
Northwest	3	1	1	-	1	-	4	-	-	1	1	2	-	-	2	4	20
Northeast	-	-	1	-	-	-	3	-	1	-	-	2	-	1	-	4	12
Southeast	2	1	1	4	-	1	3	1	-	1	-	1	-	-	-	4	19
All	8	2	6	5	1	3	12	2	1	2	1	7	1	2	2	15	70

The following table gives the average percentage all the minor tests count by districts and geographical sections.

TABLE XLIV

THE AVERAGE PERCENTAGE ALL MINOR TESTS COUNT TOWARD A STUDENT'S FINAL MARK FOR A REPORT PERIOD

District	Average Percentage*	Geographical Section	Average Percentage*
First class	25.00	Southwest	31.00
Second class	35.00	Northwest	34.00
Third class	29.00	Northeast	41.00
County High School	42.00	Southeast	27.00
All	32.00	All	32.00

*Approximate figures

Lee asked various teachers in a questionnaire what per cent the final examination counts in determining the mark of the course.[7] The median per cent was 26.1. The number of teachers answering in Lee's survey on this item was 1,152. Lee then asked the teachers what per cent do all the tests given count in determining the final grade. The median per cent was 36.0. The number of teachers answering on the above item was 1,222. Lee had originally analyzed this data according to departments, but noted very few differences, especially on the item of final ex-

7Ibid., pp. 259-262.

aminations. The mathematics, science, social science, and commercial departments put more stress on the test results for marks than did the others. Lee put a third question to the teachers, asking them how the pupil's test paper was marked in most of their self-made tests. The teachers were divided into three groups, those that had training in measurements, those that had some training in measurements, and those that had no training. The differences in results were greater between the group that had no training and the groups that had training than the differences in results between the two groups that had training. The teachers who had no training marked their papers using per cents, 100 per cent being perfect. The teachers in the two groups with training marked their papers in comparison with the other papers of the class to a much greater extent than the teachers with no training. A large number of teachers marked their papers on some hypothetical standard.

F. TEACHER-APPRAISAL OF TWO ELEMENTARY ALGEBRA PROBLEMS

Each teacher in the survey was given two elementary algebra problems with the problems already worked out. They were asked to grade these problems any way they wished on a ten-point scale. In addition each teacher was requested to explain why he marked the problem as he did.

The first problem is a fractional equation to be solved. The problem may be found in the F section of the

questionnaire in Appendix A. The "hypothetical" student in working this problem found the lowest common multiple correctly. There are three fractions in the problem to be cleared. The student did not clear the first fraction correctly, having multiplied the numerator of that fraction by the wrong factor in the lowest common multiple. The other two fractions were cleared correctly. The following multiplication was carried out correctly. However, one term was transposed without a change in sign. The "incorrect" answer was then derived. Two errors were made in the problem, one in the clearing of a fraction and the other in transposing. Some teachers in correcting this problem missed the error of transposition. However, the assumption was made that that teacher might be prone to miss errors occasionally. The following tables give the scoring on the first problem by frequency of responses for each mark assigned. The problem was marked on a ten-point scale, indicated on the top of the tables.

TABLE XLV

TEACHER-EVALUATION OF THE FIRST PROBLEM ACCORDING TO DISTRICT

District	0	1	2	3	4	5	6	6½	7	8	9	10	No Reply	Total	Average Mark
First class	3	-	-	-	-	-	-	-	-	-	-	-	1	4	0.00
Second class	10	-	4	1	2	2	1	-	2	4	1	2	-	29	3.79
Third class	14	1	1	-	-	5	4	1	-	2	-	2	3	33	3.15
County High School	3	-	-	-	-	1	-	-	-	-	-	-	-	4	1.25
All	30	1	5	1	2	8	5	1	2	6	1	4	4	70	3.17

TABLE XLVI

TEACHER-EVALUATION OF THE FIRST PROBLEM ACCORDING TO GEOGRAPHICAL SECTION

Geographical Section	0	1	2	3	4	5	6	6½	7	8	9	10	No Reply	Total	Average Mark
Southwest	9	-	2	-	1	2	-	-	-	2	-	2	1	19	3.00
Northwest	12	1	1	-	1	-	1	-	-	1	1	1	1	20	2.11
Northeast	2	-	1	1	-	3	2	1	1	1	-	-	-	12	4.46
Southeast	7	-	1	-	-	3	2	-	1	2	-	1	2	19	3.65
All	30	1	5	1	2	8	5	1	2	6	1	4	4	70	3.17

The teachers in the larger high schools tend to hold a strict marking policy. The teachers in the smaller high schools are more lenient in their policy of grading. However, the second class teachers tend to be more lenient in this respect than the third class teachers. The teachers in the County High Schools having first class enrollments graded this problem a "zero". The teachers in the eastern part of the state appeared to have a more lenient marking policy than those in the western part of the state. This might be because the western part of the state has more of the larger high schools.

The total years of experience might be a factor affecting a teacher's marking policy. The experienced teacher might be set in his ways, be more conservative, and believe that a mathematics problem is either right or wrong. His marking policy will probably be strict. The inexperienced teacher is probably more liberal and will make allowances for a wrong answer if the method is correct. Again, the experienced teacher might make more allowances for a wrong answer than the inexperienced teacher, lest he fail too many students and bring failure upon himself as a teacher. The inexperienced teacher might be strict at first believing mathematics an exact subject, but with experience he finds it more practical to be lenient. Thus, there are two sides to the picture. The following table gives the average years of experience of all the

teachers assigning each mark on the ten-point scale for the first problem. There does not seem to be much correlation between years of experience and a strict marking policy. This may be, as has been pointed out, that a teacher with so many years of experience may react either way in his marking policy. In some cases so few teachers assigned a certain mark that the calculated average years of experience for that mark may not be reliable.

TABLE XLVII

THE AVERAGE YEARS OF EXPERIENCE OF ALL TEACHERS ASSIGNING EACH MARK ON THE TEN-POINT SCALE FOR THE FIRST PROBLEM

Mark	Number of Teachers Giving that Mark	Average Years of Experience
0	30	14.03
1	1	3.00
2	5	4.60
3	1	25.00
4	2	7.50
5	8	7.63
6	5	3.80
6½	1	35.00
7	2	16.50
8	6	7.50
9	1	13.00
10	4	3.25

These teachers assigned these marks for various reasons. A teacher would assign a zero grade because work in mathematics must be accurate, because there is no excuse for carelessness, because the student shouldn't be satisfied with being half right, and the like. Teachers who assigned a higher grade to the first problem thought that effort should be worth something, that one should not grade too harshly, that the student had the right method, and the like. Some of the teachers who assigned a "zero" to the first problem stated that they did so for the following reasons:

> "If I had just covered this particular portion of the work, I would not allow credit for anything but the right answer."
>
> "I don't believe in giving credit for a problem done <u>almost</u> right as it encourages sloppy work and they become satisfied with doing them almost right."
>
> "Even though the method of solution is correct, <u>accuracy is necessary</u>. More emphasis should be placed on accuracy. Furthermore, if the student had checked his work, he would have found his answer to be in error and then checked over his work and probably could have found his error in computing the answer."

One teacher who gave a mark of "two" did so for this reason:

> "Multiplied by wrong factor, but in general it appeared that he understood the work basically. This all makes for 5 points, but the improper (or incorrect) transposing takes off 3 more points, leaving 2 points."

The teacher who gave five points to the problem gave this reason:

> "Since there are two mistakes which are easy to make even by advanced algebra students, I give half

> credit. From the way the problem is set up and worked thru, there is indication of at least some progress. By grading too harshly in lst. yr. algebra, the teacher can cause a student to lose interest in the course."

This teacher assigned eight points:

> "After finding the lowest common multiple, I think it is important for the student to realize that each numerator must be multiplied by the lowest common multiple, and that is why the denominators vanish."

Finally, the teacher who assigned ten points or perfect had this to say:

> "On paper, the steps show that the student knew his theory of operation well. (However, in speed test he could eliminate steps.)"

However, in contrast to the reason just stated, another teacher who gave five points said:

> "Should have shown more work on the paper when he cleared the fractions."

The second problem the teachers were given to grade is a story-type of problem also giving a fractional equation to solve when correctly set up. This problem may be found after the first problem in the F section of the questionnaire in Appendix A. The "hypothetical" student in working this problem did not set the problem up correctly. The resulting "wrong" equation, however, was worked correctly. The error in the second problem was different from those in the first problem. This error resulted from the misinterpretation of the conditions of the problem. The errors in the first problem were made in working with an equation already set up. The following

tables give the scoring on the second problem by frequency of responses for each mark assigned. This problem, like the first, was marked on a ten-point scale.

TABLE XLVIII

TEACHER EVALUATION OF THE SECOND PROBLEM ACCORDING TO DISTRICT

District	0	2	2½	3	4	5	6	7	8	9	10	No Reply	Total	Average Mark
First class	4	-	-	-	-	-	-	-	-	-	-	-	4	0.00
Second class	16	2	-	4	2	1	1	-	-	1	1	1	29	1.93
Third class	14	-	1	2	1	5	-	1	1	-	3	5	33	2.95
County High School	3	-	-	-	-	1	-	-	-	-	-	-	4	1.25
All	37	2	1	6	3	7	1	1	1	1	4	6	70	2.21

TABLE XLIX

TEACHER-EVALUATION OF THE SECOND PROBLEM ACCORDING TO GEOGRAPHICAL SECTION

Geographical Section	0	2	2½	3	4	5	6	7	8	9	10	No Reply	Total	Average Mark
Southwest	14	-	-	1	-	2	-	-	1	-	-	1	19	1.17
Northwest	13	-	-	1	1	2	-	-	-	1	1	1	20	1.89
Northeast	1	2	1	2	-	3	1	-	-	-	2	-	12	4.46
Southeast	9	-	-	2	2	-	-	1	-	-	1	4	19	2.07
All	37	2	1	6	3	7	1	1	1	1	4	6	70	2.21

Again, the teachers in the larger high schools hold a stricter marking policy. As the high schools become smaller, the teachers become more lenient in their marking policy. Two of the four County High Schools have first class enrollments; the other two, second class. The teacher from the County High School with a second class enrollment assigned a "five" to this problem. Again, the teachers in the eastern part of the state hold a more lenient marking policy than those in the west.

As in the case of the first problem, the experienced teacher does not seem to have a stricter marking policy than the inexperienced teacher. However, so few teachers assigned certain marks that the calculated average years of experience for those marks are probably not reliable. The following table gives the average years of experience for all the teachers assigning each mark on the ten-point scale for the second problem.

TABLE L

THE AVERAGE YEARS OF EXPERIENCE OF ALL TEACHERS ASSIGNING EACH MARK ON THE TEN-POINT SCALE FOR THE SECOND PROBLEM

Mark	Number of Teachers Giving that Mark	Average Years of Experience
0	37	13.19
2	2	2.00
2½	1	35.00
3	6	15.00
4	3	3.33
5	7	2.29
6	1	15.00
7	1	3.00
8	1	2.00
9	1	13.00
10	4	11.75

The teachers in general assigned lower marks to the second problem than to the first. Many thought that the equation must be set up correctly according to the conditions of the problem. In other words, the solution of an equation is worthless, even though it may be correctly done, if the equation itself is not correct. The teachers must have thought it more important for the student to set up an equation correctly than to work it correctly. Some teachers who assigned a "zero" to the second problem stated these reasons:

> "I give no credit for a problem having the wrong answer, even if the last step in a long solution is the only error."
>
> "I do not give any credit when the equation that the student sets up is wrong."
>
> "First, the important part of this problem is to set up the correct equation. This has not been done. Although the mechanics are OK for steps 2 and 3, the answers obtained are obviously wrong by quick inspection. Students should watch for answers which are not reasonable."

This teacher assigned a mark of "four":

> "Student deserves credit for putting unknowns down correctly."

However, another teacher who gave a mark of "zero" said:

> "Unknowns not expressed properly."

One teacher who assigned a mark of "seven" gave this reason:

> "One mistake. Method correct. Seemed to know what he was doing except for the X under the three."

This teacher assigned a mark of "ten", but might have

missed the mistake:

> "This problem involves reading, thinking, and planning the final equation. While we do teach math., I feel reading, etc. are important also. Perhaps our greatest weakness is the small number of written problems given to students. They are weak in written problems, thus we stay away from them possibly to help their mark. The fundamentals are found in this problem, but to me the equation is the most important."

The last statement made by this teacher indicates that the mistake, that of setting up the equation, might have been missed.

Trabue points out that a teacher who assigns a mark believes that mark to be very significant. Thus, a teacher who gives a pupil a certain grade would usually be unwilling to admit that the pupil deserves only a lower grade. This teacher would probably be correct by his own standards. One teacher might carry a mental scale of what degrees of scholarship deserve certain grades. Each teacher seems to have his own mental scale, which does not transfer to other teachers as such. A teacher might not agree with the scale of another teacher if such a transfer were possible. Thus, teachers' standards become subjective in nature.[8]

Orleans believes that the lack of any precise standard for evaluating answers causes teachers to vary

---

[8]Marion Rex Trabue, Measuring Results in Education. (New York: American Book Company, 1924), pp. 49-51.

considerably in the marks they give independently to the same pupil response. He points out that studies of the traditional examinations have shown under such independent marking that teachers' marks very often vary by more than fifty per cent on the same pupil's paper. Studies have also shown that the same teacher after several days varied considerably in his marks on the same papers. Thus, a teacher may even differ with himself.[9]

According to Orleans, the chief objection to the traditional test is the meaninglessness of the marks which result from the subjectivity of the marking. The marking of a test is subjective to the extent that persons who are competent to mark the answers to a question or problem will differ in the value they assign to a given answer. Orleans says that an experiment where a group of teachers would mark independently a single pupil's paper would probably show a significant variation in the results. Teachers will also vary in their reasons why they gave the mark.[10]

The standardized test usually has a definite answer to every question or problem. Every algebra problem usually has a definite answer. A standardized test

---

[9]Jacob S. Orleans, Measurement in Education. (New York: Thomas Nelson and Sons, 1937), pp. 49-42.

[10]Ibid., pp. 53-54.

will usually allow no part or fractional credit for any answer. Such a procedure could be followed in the case of an algebra problem. Teachers could easily agree on results if they marked problems on an "all or nothing" basis. Around half of the Montana teachers in the survey appear to do just that. Some of these teachers feel that accuracy is one of the prime objectives of mathematics and important enough to strive for. Theywould have their students organize their facts correctly and may feel that a little effort on the part of the pupil giving the right answer is much better than a lot of effort leading to a wrong result. Teachers allowing partial credit would emphasize factors other than accuracy. Such factors might be understanding of theroblem and effort expended by the student. The writer feels that accuracy should accompany understanding and effort to produce fruitful results in problem solving. Understanding need not be sacrificed in striving for accuracy. The "all or none" system of grading can encourage accuracy among students and produce a more consistent grading procedure among teachers.

CHAPTER IV

THE EVALUATION OF ELEMENTARY ALGEBRA TESTS SUBMITTED BY MONTANA TEACHERS OF ALGEBRA

A test may be judged on two criteria, validity and reliability. A general measure of validity can be obtained by a determination of the difficulty of these algebra tests, by a study of requirements for elementary algebra in the Montana course of study for mathematics, by a comparison between the different algebra examinations given by different teachers, and by a study of the kinds of test items given in these tests. A general measure of reliability can be obtained by a study into the objectivity, length, and evenness of scaling of these algebra tests.

A. VALIDITY

The most important criterion of a test is its validity. Greene, Jorgensen, and Gerberich define the validity of an examination as depending upon the efficiency with which the examination measures what it attempts to measure. A test must measure the trait it intends to measure in order to maintain its validity. The validity would be lowered as one deviated more from this trait or category of the test. Thus, a pupil's ability to solve written problems cannot be measured so well if linguistic difficulties are introduced. Then the text is measuring

the pupil's ability to read. However, problem solving ability and linguistic ability cannot be completely separated. The validity of tests might differ because no two teachers teach exactly the same course and that no one teacher teaches exactly the same course twice during his lifetime.[1]

Greene, Jorgensen, and Gerberich give a discussion of curricular validity. This type of validity is the most important, as any method of test validation must be based on relatively subjective judgment concerning the degree to which the test covers the proper ground. A teacher attempts to insure curricular validity in a test by making certain that the test deals with the types of educational outcomes he wishes to measure and that the test is at the proper level of difficulty for his pupils. The validity of a test may be measured by how much the test stresses the educational outcomes given in the course of study. The validity of an achievement test based more generally upon information than upon skills depends largely upon the opportunity the pupil has had to master the content in the dourse as measured in the test. In this case the teacher is probably the best judge of validity, as he knows what

---

[1]Harry A. Greene, Albert N. Jorgensen, and J. Raymond Gerberich, Measurement and Evaluation in the Secondary School. (New York: Longmans, Green and Company, 1943), pp. 52-73.

he has taught.[2]

Jordan says a test is valid in proportion as it measures what it purports to measure. Validity cannot be defined by the knowledge of what a test measures, since it is known that a test does not satisfy its claims. Thus, validity is better defined by what a test attempts to measure.[3]

Jordan gives a discussion of internal validity. The internal validity refers to the care with which the items of the test are selected and arranged. The items of the test are constructed after a consideration of the desired objectives. Such items are carefully written, judged by a jury of experts, and then tried out on a small sample of subjects. Ambiguous statements appearing must be modified or omitted entirely. Further revisions might be necessary before the test can assume its final form. A teacher can follow these procedures in giving tests, eliminating poor items and keeping or modifying the better items. Jordan feels that the teacher would be in the best position to make the most valid test for his class. Thus, an elementary algebra teacher would know the areas he had taught and the objectives he had had in mind.

---

[2]Loc. cit.

[3]A. M. Jordan, Measurement in Education. (New York: McGraw-Hill Book Company, Inc., 1953), pp. 14-21.

The teacher gives a test over specific subject matter in any area. The standard test is different from the teacher-made test in this respect. A standardized test would base its items over any area on subject matter common to courses of study and popular textbooks. The standardized test-maker would select items according to frequency of occurrence. Such a system tends to neglect local materials introduced for interest and to retain common facts in the test. Thus, as has been pointed out, a standardized test may be lacking in social utility. The teacher must be careful that he does not "over-standardize" his tests to this extent. Such a test would be concerned with the mere reproduction of facts. Facts would not be presented in new situations, and students would be prone to memorize specific facts for the sake of memorization. Desirable educational objectives must be dominated by social utility, which is always reflected in a valid test.[4]

McCall lists various criteria which contribute to the validity of tests. The first of these criteria is that tests should be in harmony with the philosophy of education. The effort of the student can be easily misdirected by examinations out of harmony with the philosophy of education.

Next, tests should measure organization of memoriter

---

[4]Loc. cit.

learnings. Tests should not emphasize mere memoriter learnings as the student might become satisfied with just memorizing. A test item which requires some application or integration of information can encourage better habits of study. Thus, statement problems in elementary algebra might make better items than the so-called drill problems for the organization of knowledge acquired. The student learns the fundamental processes of algebra through these drill problems. An excessive use of this kind of problem might lead the student toward mere memoriter learning. Statement problems should be introduced when the drill problems have been mastered to enable the student to organize the knowledge he has acquired.

A test should also give a good supply of diagnostic information. In diagnostic tests, a pupil's performance on a certain item or group of items is of more importance than his total score.[5]

McCall feels that the completion test is the least useful as an achievement test. The completion test tends to encourage verbal learnings, may have too many words removed, is easily mutilated at the wrong place, and may be difficult to score. The essay test should only be used to test a pupil's ability to organize a complicated subject. The matching type is seldom used because it cannot be

---

[5]William A. McCall, Measurement. (New York: The Macmillan Company, 1939), pp. 29-54.

easily employed on as many kinds of subject matter as the multiple-choice and because the matching type is not as popular with pupils. McCall feels that the alternate-choice and multiple-choice are the most useful and popular of all types. The right choice in a multiple-choice test should not appear obvious and its position among the other items should be determined only by chance.[6]

The simple-recall type and not the multiple-choice type is the best to use when testing recall. The simple-recall might also be used when there is just one possible brief answer, when objective scoring is not absolutely necessary, and when electric machine scoring is not planned. Algebra problems present the simple-recall type. The student works the problem and usually arrives at one brief answer. Fewer items would be required to make the test reliable as there is little or no guessing by the pupils. The simple-recall is probably the best type to use with algebra problems. No answer is suggested and the student must work the problem to aid in arriving at the answer. The teacher should require the presence of the student's work in a logical algebraic manner along with the correct answer. Such a procedure would prevent the student from using processes other than algebra to arrive at the answer.

---

[6] Loc. cit.

Multiple-choice tests may have from three to about five or six items to choose from McCall states that the use of five or six choices seems to be somewhat superior to the use of three or four.[7] If such a form is used in an algebra test in connection with problems, the student should be required to work the problem first in a logical, orderly manner.

Objective tests may be lacking in comprehensiveness. A comprehensive measure of a trait is a criterion of validity. Comprehensiveness may only be feasible when a narrow ability or a limited field of subject matter is being measured. An objective test may be made more comprehensive by the inclusion of random samplings of the trait in question. Enough of these random samplings should be used to give a reliable measure of the pupil. The method of random sampling makes possible the construction of alternate tests. This method, however, only tells what per cent of a total field of knowledge the pupil knows. Comprehensiveness may be secured in this respect, especially with skill tests, by the use of type material. Each subject involving skill is assumed to contain typical units or typical processes. Thus, ability in the trait or subject is generally determined by measuring ability in the type processes.

---

[7] Loc. cit.

The test should correspond somewhat to life situations. Some tests correspond more to life situations than other tests do. Many algebra tests present a series of isolated drill problems showing little or no relation to life situations. Such problems are necessary for testing the pupil's ability to perform the fundamental operations. However, statement problems presenting practical situations could be emphasized more than they are. Teachers probably use these items to a minimum to avoid making the test too "hard" for their pupils.

Lee states that although most objective type questions are slightly more valid than the essay question, the true-false question is slightly less valid than the essay.[8] The various types of objective questions have been found to be of almost equal validity with the completion type slightly more valid than others. McCall had stated that the completion test is the least useful as an achievement test and the true-false one of the most useful and popular of all types.

Individual test items may be validated in a simple manner by comparing the responses of the best third or fourth of the pupils with those of the poorest third or fourth. Lee cites a study by Kelley showing that the upper and lower twenty-seven per cent of the pupils give

---

[8]J. Murray Lee, A Guide to Measurement in Secondary Schools. (New York: D. Appleton-Century Company, Inc., 1936), pp. 331-334.

the best contrast. The test papers of the upper and lower twenty-seven per cent are selected. The number of pupils getting each question right is tabulated for each group. The questions or problems which show the largest amount of differentiation between the groups in favor of the best group are the most valid questions or problems.[9]

Validity is related to the selection of test items. A test should not include too wide a range of topics. Monroe gives an example of a mathematics test having six different topics included in the ten problems. Thus, a test may attempt to measure too many things and probably will not measure any one thing well as a result. Even if the test scores were accurate measures of what the pupils could do on the test, they would not have a definite meaning. The score would give no evidence as to what topic the pupil was lacking in. The test items in general should have some relation to each other. Otherwise, separate scores should be tallied for topics completely different from each other.[10]

---

[9] Loc. cit.

[10] Walter Scott Monroe, Measuring the Results of Teaching. (Boston: Houghton Mifflin Company, 1918), pp. 17-18.

B. RELIABILITY

A test functioning consistently is said to be reliable. Greene, Jorgensen, and Gerberich define the reliability of an examination as depending upon the efficiency with which a test measures what it does measure. A test may measure well what it does measure, but poorly what its user attempted to measure. A test must measure well what it does measure before it can measure with success what it attempts to measure. Test reliability may exist without validity, but a valid test must also be reliable. Reliability may be thought of as a phase of validity.

Reliability might be expressed by the coefficient of correlation. The coefficient is obtained by correlating scores on two equivalent forms of the same test given in order by the same procedure to the same group of pupils. Such a measure is the reliability coefficient or coefficient of reliability. Both the reliability and validity coefficients are special applications of the coefficient of correlation.

Tests do not attempt to measure every outcome of instruction. A more practical method, that of a system of sampling, is employed. Test items can be constructed representing the types of pupil outcomes expected. The scores resulting from such tests may be accepted as representative of the relative achievement of the pupils for the

whole area sampled by the test items. Adequacy of sampling is a phase of reliability. A small sampling of material may cause a pupil's score to be raised or lowered from what it should be. The student may or may not be acquainted with the limited material. An adequate sampling of material must be taken if the test is to be fair for all concerned and if the test is to be reliable.

The development of the objective test arose out of the need for the elimination of the subjective factor in the marking of examination papers. Test objectivity aids in the elimination of the opinion, bias, or judgment of the scorer. The objective test items are usually so stated that only one answer suffices. Many mathematical problems come under this category.

The simple-recall type is one form of the objective exercise. The solution of algebra problems generally represents this type of response. Other types of response sometimes used in algebra tests are the alternate-choice, multiple-choice, matching, and completion.

Test objectivity should make for perfect agreement as to what is the correct answer. However, such objectivity becomes meaningless when teachers are willing to accept incorrect answers. A considerable number of Montana teachers of elementary algebra corrected the problems presented in the questionnaire allowing partial credit. The disagreement among these teachers represents the presence of subjective

factors in the scoring. This situation calls for a return to the evils of the traditional examination. Only the correct answer should be accepted in an algebra examination. In theory, students are not being taught how to do problems incorrectly. A test really has no objectivity at all when the scorer loses respect for the correct response. Test objectivity is another phase of reliability.

Reliability depends upon the ease and accuracy with which the test can be administered. This phase of reliability is termed "administrability". Ease of test administration depends upon clear, simple, and direct instructions to the pupils. Samples should be used to indicate any particular form of response desired. The nature of any response required should be within the range of pupil ability. Long and involved instructions to the pupils are best avoided. The test itself should not ordinarily require elaborate preparatory arrangements on the part of the teacher. The appearance of the test, page size, length of line, and the size of type affect test administrability.

Test results should be obtained in as simple, rapid, and routine a manner as possible. Tests containing only one possible answer per item should be easy to score. The system of partial credit introduces added considerations of weighing and evaluation on the part of the teacher. Ease of scoring is lost and the test lacks in scorability, another phase of reliability.

A test has utility to the degree that it serves a

definite need satisfactorily in the situation in which it is used. A test should have definite purposes with its results used in an intelligent attempt to satisfy the needs at hand. The teacher must employ foresight in constructing a usable test and must be able to use the results in serving the needs of the classroom.[11]

Jordan defines reliability as implying precision or accuracy.[12] In validity the emphasis is on a test's agreement with the objective; in reliability, upon agreement with itself.

Jordan goes on to say that various factors affect reliability. A pupil will respond to a test according to interest and effort, physical conditions, emotions, and thought processes already in progress. Variation in pupil response would lower the reliability of a test, even if the test were a perfect measuring instrument and the conditions under which the test was given were ideal in every way.

The teacher himself may be a factor in lowering the reliability of atest. He may not give clear instructions to his pupils for taking the test, may not allow enough time for the test, may be unaware of the lack of independent work on the part of his pupils, and may not give his

---

[11]Greene, Jorgensen, and Gerberich, loc. cit.

[12]Op. cit., pp. 26-33.

pupils enough encouragement to do their best with the course.

The teacher may be inaccurate in the scoring process. This situation was shown when some of the Montana teachers of algebra would miss the errors presented in the two problems on the questionnaire. The allowance of partial credit, as was brought out, would detract from the reliability.

Other factors, such as good pupil motivation throughout the test, emotional calmness, careful test administration, and effective scoring, can help to raise the reliability of a test. A longer test with adequacy of sampling is a more reliable test.

The confidence that can be put in the individual pupil score is questionable. The best interpretation of the score probably comes out of the concept of the variation of the obtained score. Jordan cites a formula for the calculation of the variability of the obtained score.[13] This variability is equal to the standard deviation of the test multiplied by the square root of one minus the reliability of the test. This formula might be expressed as

$$V = S.D. \left(1 - r \right)^{\frac{1}{2}}$$

Thus, the larger the standard deviation and the smaller

---

[13] Loc. cit.

the reliability, the more the individual test score will vary. If the variability were four and the test score were sixty, the chances would be better than two in three that the true score lies between fifty-six and sixty-four. The smaller the variability, the more confidence can be put in the score made on the test.

McCall defines reliability as the amount of agreement between results secured from two or more applications of a test to the same pupils by the same examiner. Perfect reliability would be obtained when the same examiner gives two identical tests according to an identical procedure to the same pupils. The absence of these factors cause unreliability in tests.[14]

The coefficient of reliability varies between 1.00 and - 1.00. A coefficient of 1.00 indicates a perfect reliability and the test is completely objective. Less reliable tests have coefficients of less than 1.00. A test with a coefficient of 0.00 has no reliability and measures in a haphazzard manner. The test results may or may not be significant. Prediction through such a test is impossible. A coefficient of -1.00 indicates a perfect negative reliability. In such a case two equivalent forms of the same test would give opposite results. The more intelligent students can be steered toward incorrect responses which appear obvious. The less intelligent

---

[14]McCall, op. cit., pp. 55-60.

students may not be able to see the incorrect response as obvious and are at an advantage with this item. More of the less able students may answer the item correctly, and the item has a negative value as a result. Too many of these items in a test may lower the reliability to practically nothing or even to a negative coefficient.

Lee states that numerous studies have shown that objective tests have much higher reliability than essay tests. Special care must be used in evaluating essay tests if their reliability is to be raised.[15] Some algebra tests do have items where the student is to define or clarify. Self-expression is demanded on the part of the student when he deals with these "essay-type" items. Thus, the danger of subjective evaluation on the part of the teacher enters in. The danger could be greater when the teacher allows partial credit or marks subjectively on the objective items. The teacher should evaluate these subjective items with care or avoid their use altogether.

Lee feels that the recall type of item is probably the most reliable, followed by the recognition types.[16] Lee feels that the true-false question, as well as being the least valid, is also the least reliable. The reliability of the alternate-choice test can be raised by modifications. Such modifications as requiring the student

---

[15]Lee, op. cit., pp. 336-339.

[16]Loc. cit.

to correct the wrong statements would raise the validity of the test and the reliability as well.

A test having many items is more reliable than the test with few items. However, the more reliable the type of item is, the fewer the items of that type would be needed to maintain the reliability of the test. A greater number of less reliable items would be needed in the test to maintain the reliability.

C. AN EVALUATION OF THE ELEMENTARY ALGEBRA TESTS OF MONTANA TEACHERS AS RESPECTING VALIDITY AND RELIABILITY

Each teacher to whom a questionnaire was sent was requested to enclose a copy of any major elementary algebra test and the pupil scores on that test. Forty-two teachers returned tests. However, as some of the tests were minor tests, thirty-six major tests are considered in this part of the chapter. Nineteen of these are semester and final tests and seventeen are six weeks' tests. Twenty-one teachers returned pupil scores. However, some of these scores pertained to minor tests. Consequently, only fourteen sets of pupil scores will be considered.

In order to become better acquainted with these tests and to gain some idea of their length and difficulty, the writer administered to himself all thirty-six of these algebra tests. The time in minutes spent on each test was divided by the number of minutes the teacher allowed for the test as stipulated on the questionnaire.

All the decimal quantities obtained for the semester and final tests and the six weeks' tests were added together and averaged. Thus, if the writer spent twenty minutes on a test where the teacher allowed one hour, the decimal quantity obtained would be 20/60 or 0.33. An average of 0.50 was obtained for the decimal quantities of the semester and final tests. Thus, the writer did these tests on the average of half the time allowed by the teacher. The typical elementary algebra student would then take twice the time to do the test as the writer or forty minutes if the writer took twenty. Thus, if sixty minutes were allowed for the test, the test is not long enough or is too easy. An average of 0.49 was obtained for the decimal quantities of the six weeks' tests. The writer did these tests on the average of less than half the time allowed by the teacher. Then, the typical elementary algebra student would take slightly more than twice the time of the writer to do the test. The following table gives the decimal quantities for the semester and final tests and the six weeks' tests.

TABLE LI

DECIMAL QUANTITIES OF MAJOR ELEMENTARY ALGEBRA TESTS AND THEIR RANK

Semester and Final Tests		Six Weeks' Tests	
Rank	Decimal Quantity	Rank	Decimal Quantity
1	0.84	1	0.77
2	.83	2	.73
3	.80	3	.70
4	.66	4	.68
5	.65	5½	.60
6	.62	5½	.60
7	.59	7	.56
8½	.50	8	.53
8½	.50	9	.52
10	.49	10	.45
11	.48	11	.43
12	.44	12	.39
13½	.37	13	.36
13½	.37	14	.33
15	.34	15½	.27
16	.32	15½	.27
17½	.27	17	.13
17½	.27	Average	.49
19	.17		
Average	.50		

The typical or average algebra student probably would not finish a test having a decimal quantity appreciably greater than the average in the time allowed. The test having suitable length would have a decimal quantity around the average.

This procedure of computing test length is not completely reliable as the factors of difficulty and test length, both of which can make a test long, were not separated. However, a reasonable idea of test length should be gathered from this procedure.

Standard deviations were computed for the fourteen sets of pupil scores. The standard deviation or sigma ( ) is a measure of the variation of the scores from the mean. Test scores having a very large range, the difference between the highest and lowest scores, would have a large standard deviation. The formula for the standard deviation may be expressed as

$$S.D. = \left[Ms^2 - (Ms)^2\right]^{\frac{1}{2}}$$

Each score is squared and then totaled. The mean obtained from the squared scores is represented by Ms^2. The mean is obtained for the scores and then squared. The squared mean of the scores is shown as $(Ms)^2$. The mean of the squared scores is either greater than or equal to the squared mean. If allthe scores of a test were identical, the mean of the squared scores would be

equal to the squared mean. Then there would be no standard deviation as there is no variability in the scores.

The formula for the variability of the individual score,

$$V = S.D. (1 - r)^{\frac{1}{2}}$$

was brought out earlier in this chapter. If the test has a perfect positive reliability (1.00), the variability of the individual score would be zero and the standard deviation would have no effect. The standard deviation would lower the reliability of the individual score to a greater degree as the test becomes less reliable.

A summary of the variability of the pupil scores is presented in the "C" section of the next chapter.

Validity: All of the algebra tests were placed into four categories by the writer - excellent, good, fair, and poor. The following gives a summary of these tests.

Nine out of the thirty-six tests in the survey were rated as excellent by the writer. All of these tests are satisfactory on the three most important points of validity; namely, level of difficulty, organization of memoriter learnings, and comprehensiveness. These three items of validity would also relate to "understanding" on the student's part. The next three points of validity in order of importance are social utility, proper limiting of types measured, and avoidance of linguistic difficulties.

A first semester examination in this group contains these kinds of items:

TABLE LII

TYPES OF ITEMS IN A SEMESTER TEST
RATED AS EXCELLENT AND THEIR FREQUENCY

Recall Types	Frequency
Substitution	1
Combination of Addition, Subtraction, Multiplication and Division	1
Linear Equations	4
Addition	3
Subtraction	2
Multiplication	6
Division	3
Removal of Parentheses	1
Factoring	7
Statement Problems	12
Total	40

The level of difficulty of this examination is moderate, not too hard and not too easy. The statement problems present applications of linear equations, decimal linear equations, and formulas. These statement problems aid in the organization of memoriter learnings. Twenty-seven out of the forty problems are comprehensive. The items on linear equations, addition, subtraction, division, and removal of parentheses lacked in this respect. The test has a moderate amount of social utility, does not attempt to measure too many things, and presents no undue linguistic difficulties. A copy of this semester examination may be found in the appendix.

An algebra six weeks' test in this group contains these kinds of items:

TABLE LIII

TYPES OF ITEMS IN A SIX WEEKS' TEST
RATED AS EXCELLENT AND THEIR FREQUENCY

Recall Types	Frequency
Division	7
Linear Equations	8
Statement Problems	7
Total	22

The writer considered this test to be of moderate difficulty. The statement problems aid in the organization of memoriter learnings. These items present applications of

linear equations, one of which is fractional. One statement problem presents expression in algebraic terms. All of the types of problems presented are comprehensive. The test has a moderate degree of social utility, does not attempt to measure too many things, and presents no unnecessary linguistic difficulties. A copy of this six weeks' test may be found in the appendix.

Five out of the nine tests in this group were considered satisfactory on all of the six points of validity by the writer. The other four tests were found to be satisfactory on five of the six points. All four are lacking only in social utility.

Eighteen other tests were rated as good by the writer. All of the tests in this group are satisfactory on two of the three most important points of validity.

A first semester test in this group contains these kinds of items:

TABLE LIV

TYPES OF ITEMS IN A SEMESTER TEST
RATED AS GOOD AND THEIR FREQUENCY

Recall Types	Frequency
Addition	25
Subtraction	25
Multiplication	25
Division	25
Linear Equations	50
Total	150

This semester test is of moderate difficulty. However, the test does not seek to organize memoriter learnings. The inclusion of some application problems might have helped to remedy this. The test as a whole is comprehensive enough to give an adequate measure of ability with each type of item. The test lacking in problems of a practical nature does not show any social utility. However, the test does not attempt to measure too many things and does not present undue linguistic difficulties. A copy of this semester test may be found in the appendix.

An algebra first six weeks' test in this group contains these kinds of items:

TABLE LV

TYPES OF ITEMS IN A SIX WEEKS' TEST RATED AS GOOD AND THEIR FREQUENCY

Recall Types	Frequency
Formulas	1
Addition	7
Expression in Algebraic Terms	8
Factoring	1
Completion	4
Substitution	5
Combination of Addition and Division	1
Linear Equations	3
Statement Problems	2

Recognition Types	Frequency
Multiple-choice, three items	1
Total	33

The writer considered this six weeks' test to be of moderate difficulty. The examination is weak in problems calling for applications of what the student has learned. The majority of the items should give a comprehensive measure of the ability they attempt to measure. The deficiency of statement problems has made the test weak in social utility. The test itself attempts to

measure too many kinds of items. However, no undue linguistic difficulties are introduced. A copy of this six weeks' test may be found in the appendix.

Most of the tests in this group are satisfactory on four out of the six points of validity. Most of the tests here were found to be deficient on organization of memoriter learnings and social utility. A summary of results is given in the following table. The points of validity are listed in order of importance.

TABLE LVI

A RATING OF SIX POINTS OF VALIDITY AS SATISFACTORY OR UNSATISFACTORY ON THE TESTS CONSIDERED "GOOD"

Points of Validity	Number Satisfactory	Number Unsatisfactory
Level of difficulty	12	6
Organization of memoriter learnings	6	12
Comprehensiveness	18	-
Social utility	6	12
Proper limiting of types measured	14	4
Avoidance of linguistic difficulties	18	-

The writer considered three tests out of the thirty-six to be only fair. All of the tests in this group are satisfactory on only one of the three most important points of validity.

An algebra semester test in this group contains these kinds of items:

TABLE LVII

TYPES OF ITEMS IN A SEMESTER TEST RATED AS FAIR AND THEIR FREQUENCY

Recall Types	Frequency
Simultaneous Equations	3
Statement Problems	1
Factoring	3
Quadratic Equations	5
Division of Fractions	3
Linear Equations	7
Radicals	13
Total	35

The level of difficulty for this semester examination is moderate. The test presents only one application problem and makes no further attempt to organize memoriter learnings. The test items are not as comprehensive as they should be. The examination in general is weak in social utility and would not help the student to see the practical value of algebra. However, the instructor properly limited the types of items measured. No undue linguistic difficulties are presented in the test. A copy of this test may be found in the appendix.

A second six weeks' algebra quiz contains these kinds of items:

TABLE LVIII

TYPES OF ITEMS IN A SIX WEEKS' QUIZ RATED AS FAIR AND THEIR FREQUENCY

Recall Types	Frequency
Statement Problems	2
Substitution	1
Addition	2
Subtraction	2
Removal of Parentheses	1
Linear Equations	1
Multiplication	3
Total	12

This quiz or test has twelve problems. The student could do any ten of these including three specified problems. Two of the ten problems the student has to do have two parts. Thus there are twelve items in all. The writer considered this test to be too difficult. The quiz is weak in problems of application. Even though there are four statement problems in all with only one required, the typical student would try to get away from these items as much as possible where he has the choice. The two statement problems the typical student might do have the same answers and are essentially the same with only a difference in

wording. However, all of the types of items with the exception of the statement problems are comprehensive. Two of the three items of multiplication were quite difficult in that they have more than one term in their exponents. One of these difficult items is required. Such items are probably beyond the ability level of the elementary algebra student and should not have been included in the test. The examination being weak in problems of application is also weak in social utility. However, the instructor does not attempt to measure too many kinds of items. No undue linguistic difficulties are presented in the test. A copy of this test may be found in the appendix.

A summary of results for the tests considered fair is given in the following table. The points of validity are listed in order of importance.

TABLE LIX

A RATING OF SIX POINTS OF VALIDITY
AS SATISFACTORY OR UNSATISFACTORY ON THE TESTS
CONSIDERED "FAIR"

Points of Validity	Number Satisfactory	Number Unsatisfactory
Level of difficulty	1	2
Organization of memoriter learnings	1	2
Comprehensiveness	1	2
Social Utility	-	3
Proper limiting of types measured	2	1
Avoidance of linguistic difficulties	3	-

The six remaining tests were considered by the writer to be poor. All of the tests in this group are not satisfactory on any of the three most important points of validity.

A semester examination in this group contains these kinds of items:

TABLE LX

TYPES OF ITEMS IN A SEMESTER TEST RATED AS POOR AND THEIR FREQUENCY

Recall Types	Frequency
Exponents	6
Statement Problems	5
Addition	4
Reduction of Fractions	1
Subtraction	5
Multiplication	4
Division	6
Removal of Parentheses	4
Recognition Types	**Frequency**
Multiple-choice, three items	3
Total	38

The writer considered this examination to be too easy. The statement problems involve applications of linear equations, one of which is fractional. However, these problems are too easy and would not require much organization of memoriter learnings. More statement problems could have been presented. All of the types presented lack in comprehensiveness, with the possible exception of the problems on addition and the removal of parentheses. This would leave thirty out of thirty-eight items not comprehensive. Social utility is lacking in this examination. However, the test does not attempt to measure too many things and presents no undue linguistic difficulties. A copy of this test may be found in the appendix.

A fifth six weeks' test in this group contains fourteen items in all. The student had the choice to do any ten. The following table shows the kinds of items the typical student might choose to do:

TABLE LXI

TYPES OF ITEMS IN A SIX WEEKS' TEST RATED AS POOR AND THEIR FREQUENCY

Recall Types	Frequency
Simultaneous Equations	1
Multiplication	2
Factoring	6
Linear Equations	1
Total	10

This test is considered by the writer to be too easy. The instructor included two problems of application, but the typical student having his choice of items to do probably would not attempt these more difficult problems. The test is basically poor in the organization of memoriter learnings as a result. The examination presents a lack of comprehensiveness. Social utility is lacking as the problems of application are really optional. Only one of the two problems of application would involve some thought on the part of the student. The other problem is quite simple. However, the test does not attempt to measure too many things and presents no undue linguistic difficulties. A copy of this six weeks' test may be found in the appendix.

All six of these tests were considered by the writer to be unsatisfactory in level of difficulty, organization of memoriter learnings, comprehensiveness, and social utility. However all six of these tests are satisfactory in that they properly limit the types measured and avoid linguistic difficulties. However, these last two points of validity are the more minor ones.

The following table gives an overall tabulation of the six points of validity principally discussed in the preceding:

TABLE LXII

A RATING OF SIX POINTS OF VALIDITY
AS SATISFACTORY OR UNSATISFACTORY ON ALL THE TESTS

Points of Validity	Number of Semester and Final Tests		Number of Six Weeks' Tests	
	Satis-factory	Unsatis-factory	Satis-factory	Unsatis-factory
Level of difficulty	9	10	13	4
Organization of memoriter learnings	9	10	7	10
Comprehensiveness	13	6	15	2
Social Utility	6	13	5	12
Proper limiting of types measured	15	4	16	1
Avoidance of linguistic difficulties	19	-	17	-

The following table gives a tabulation of the degrees of difficulty of all the tests considered:

TABLE LXIII

THE DEGREE OF DIFFICULTY BY NUMBER OF TESTS

Kind of Test	Difficult	Moderate	Easy
Semester and Final	3	9	7
Six Weeks'	2	13	2
Total	5	22	9

Reliability: Of the six points of reliability, the most important one is objectivity in the scoring policy of the teacher. Most algebra tests and mathematics tests in general are already objective in form. The lack of basic objectivity in a mathematics test is usually not a problem. The main problem is how the teacher treats what the student has done. Nine teachers marked their tests in the semester and final category satisfactorily and nine did not mark their tests objectively. The scoring policy of the other teacher was not made clear. Five teachers marked their tests in the six weeks' category objectively and eleven more were not objective in this respect. The scoring policy of one teacher was not clear.

The most objective kind of a scoring policy is a uniform and consistent one. Teachers who would give partial credit to the problems they mark are making their own in-

dividual judgments on the merits of the students' work. They may decide that the incorrect work of a student is worth so many points. A policy of partial credit would be justified if the number of points the teacher assigned to the students' work were the exact amount with no deviation or error; if the teacher marked all problems with exactly the same degree of judgment, and if all teachers agreed exactly how all problems should be marked. These goals are too much to strive for. Individuals will differ in their judgments, and subjectivity enters into the scoring procedure as a result. The basic objectivity of a test has little meaning if the results are interpreted subjectively. Tests decreased in objectivity are also decreased in reliability. Tests without reliability can have no validity.

A teacher should accept only the correct answer arrived at in a logical manner. A teacher would agree with himself more consistently under such a system, and many more teachers would agree among themselves. The "all or none" system of grading is probably the most objective there is. Argument will arise as to what constitutes a logical manner of the student in arriving at the correct answer. Subjectivity will enter in here. However, the teacher must require a logical procedure on the part of the students in solving problems. Otherwise, the student may arrive at the correct answer in a non-algebraic manner or by pure guess.

The teacher may use the results of a test for diagnostic purposes, for guidance, for determining period and final marks, and the like. The results of a test may or may not be usable for the purposes of the teacher. The usability of the test results will depend on the objectivity of the test, and in this case, on the objectivity of the scoring procedure. Tests marked objectively will generally produce usable results. Eight semester and final tests were considered by the writer as producing usable results and ten were not. The usability of one test could not be determined as the scoring policy of the instructor was not clear. Six tests in the six weeks' category would probably produce usable results and ten would not. The usability of one test in this category could not be determined as the instructor's scoring policy was not made clear.

The items of each type in a test should be adequately sampled. A student may be able to do some items within a type but not others. A teacher selecting a few items for a type may or may not select the problems a student can do with success. Thus, a student may be lucky with the test items and receive a mark higher than he deserves or he may not be so fortunate and fail on the items when he really deserves a higher mark. Test items should be thoroughly sampled to aid in preventing the above possibilities. The semester and final tests were poor in this respect with only three out of nineteen having a majority of

types containing items adequately sampled. The semester and final tests must each contain more types of items than the six weeks' test as they cover more material than the six weeks' test. The teacher probably fears that the test will be too long if each type has items sampled adequately. It is better to have the test long with the items sampled adequately. More time should be allowed for the test if needed. The picture is not as bad for the six weeks' tests. Eight out of seventeen of these tests had a majority of types with items adequately sampled.

Six out of nineteen semester and final tests were satisfactory in the length of test in that the student could finish it in about the time allowed. From the other thirteen, six were too long for the time allowed and seven were too short for the time allowed. Six out of seventeen six weeks' tests had satisfactory length in that the student could finish it in about the time allowed. From the other eleven, six were too long for the time allowed and five were too short for the time allowed. The writer considered a test having a decimal quantity (explained in Chapter IV) of more than plus or minus 0.10 from the average decimal quantity of its category, e. g., semester and final, to be of unsatisfactory length for the time allowed.

If the examination is too long for the time allowed, the average student might not finish, and a complete measurement of what he could do would not be obtained. If the ex-

amination is too short for the time allowed, too many students may finish early and be a disturbance to those still on the test unless they are given something else to do or are excused. The examination should be long enough that the average student will be kept busy on it throughout the test period, but that he willfinish it on time.

A test may be long enough for the time allowed, but may still be too short because of inadequacy of item sampling or may be too easy. A long test with moderate difficulty and adequacy of sampling of items is most desirable from the standpoint of validity and reliability. More time should be allowed for such a test if necessary.

All nineteen semester and final tests and all seventeen six weeks' tests were objective in form. A mathematics test generally does not lack in objectivity. The real problem is whether the scoring procedure of the teacher is objective or not.

A point of reliability of lesser importance than the others is the administrability of a test. Seventeen out of nineteen semester and final tests were considered administrable by the writer. Sixteen out of seventeen six weeks' tests were satisfactory on this score. The administrability of a test depends on the clarity of the directions to the student, the form in which the test is presented, the condition of the form on which the test is presented, and on the amount of independent work during a test by the student. Most of the tests in the survey are on mimeographed sheets

and are readable. The test would not proceed in an orderly manner if the students did not work independently.

CHAPTER V

SUMMARY

A. A SUMMARY OF THE PREPARATION, PRESENTATION, AND EVALUATION PROCEDURES OF TEACHER-MADE TESTS AS PRESENTED IN CHAPTER III

It has been estimated that the approximately 700,000 teachers from the elementary school to the college level spend forty-two million hours of time annually in preparing and scoring tests and in analyzing and interpreting the results.[1] This is based on three hours per test and 20 tests per teacher, in testing. Thus, teachers spend more than a week of school time annually in work with tests.

Preparation of Tests: Generally, elementary algebra teachers give six or eight major tests in a school year. This might include two six weeks' tests and a final per semester or three six weeks' tests and a final each semester. Most teachers give from one to two minor tests a week. Perhaps it would be more desirable to give more quizzes than that each week.

---

[1] William A. McCall, Measurement. (New York: The Macmillan Company, 1939), p. 29.

A vast majority of the teachers informed their students that a major test was coming. About half of the teachers informed their students of coming minor tests. Students should be informed of infrequent tests such as the major examinations. However, it might be better not to announce frequent tests such as those minor. Students might prepare their lessons better if they feel a test could occur.

A majority of the teachers selected their major test items from the basal textbooks used. A greater percentage of teachers selected minor test items from the basal textbooks than for those major. The student has access to the basal textbook. The teacher should not have a problem here with the student memorizing answers if he requires from the student in a test the correct work also.

Most teachers spend an hour or more preparing a major test. Over half of the teachers spend fifteen minutes or less in the preparation of a minor test. A teacher may spend much time in preparing a test, select a poor quality of items, and thus have a poor test. Time and consideration must be given to item selection if the examination is to be reliable and valid.

Administration of Tests: Around half of the teachers allowed their students fifteen minutes or less to take a minor test. A test should be long enough that the average student can finish in the time allowed. A summary of the length of major tests is taken up in the B section

of this chapter.

The majority of teachers with more than one section gave the same minor test to each section. A greater percentage of teachers with more than one section gave the same major test to each section. Teachers should at least give different minor tests to each of their sections. Minor tests are shorter than major and easier for students to communicate to each other concerning the nature of the material.

A vast majority of the teachers did not prepare alternate major or minor tests for use within a classor section, such that each alternately seated student would be taking the same test. More teachers attempted to solve problems of independent work by the assignment of alternate seats during a test. The majority of teachers assigned alternate seats during a major test, while about a third gave alternate seats for a minor test. The assignment of alternate seats is more practical than the work of preparing alternate tests and should be just as effective.

The use of mimeographed sheets is the favored form for giving major tests, while the blackboard is favored for the minor tests. The blackboard is easier to use than the preparation of mimeographed sheets. However, the teachers' writing on the blackboard should be readable and the material presented in an orderly manner. Blackboard glare should be prevented. The student has the test right before him if he takes it from a mimeographed sheet. He can tend

to the work without looking up at the blackboard ever so often. The material on the mimeographed sheet must also be readable and presented in an orderly manner.

Most of the teachers did not feel that independent work during a test presented a problem. The teachers took three approaches toward the problem of independent work, that of being lenient, moderate, or strict. The majority of the teachers, who replied to this questionnaire item, took a moderate approach. More teachers took a strict approach to the problem than alenient one. A teacher should not be completely deaf to the problem, nor should he become fanatic over it to the point of constantly threatening and punishing upon which the student might lose respect for authority. A teacher should use various measures such as alternate seats or tests to handle the problem without directly implying that he does not trust the students.

The majority of teachers placed little to no emphasis at all on the memorization of material in the testing program. The understanding of the material is more important than the rote memorization of the subject matter. Some teachers did require that the quadratic formula be memorized.

Evaluation of Tests: Practically all of the teachers do not allow the students to correct their own major test papers. At least three-fourths of the teachers do not allow students to correct their own minor test

papers. Nearly all of the teachers do not allow students to correct major test papers not their own. However, more than a third of the teachers do permit students to correct minor test papers not their own. A few more teachers would sometimes let students correct minor test papers.not their own.

A test does have instructional value when the student is allowed to correct his own paper. However, the student might take unfair advantage of such a situation. The extra answer sheet on which the student duplicates his answers was suggested in Chapter III. However, this plan requires more class time to give and correct the test.

The majority of teachers counts each major test twenty-five or thirty-three and one-third per cent toward a student's final mark for a report period. The percentage all minor tests count toward a student's final mark for a report period varied from less than ten per cent to as much as seventy-five per cent.

Perhaps the student's grade for a report period should be based entirely on tests. Lee feels that the teacher should not mark on daily work, which may not be the pupil's own work, or on the recitation, where personality, appearance, command of English, and behavior bear much influence.[2] However, the daily work could be used as an

---

[2]J. Murray Lee, A Guide to Measurement in Secondary Schools. (New York: D. Appleton-Century Company, Inc., 1936), pp. 229-231.

instructional device. The student could correct his own daily work in class and see where he made his errors. No marks should be taken on the daily work though. A quiz given to the class would best supply the mark for the day. If the daily work would provide the instructional value to the student, the test would not be needed as an instructional device. The teacher can correct all the tests, which would remain as a measurement device.

Two elementary algebra problems were presented in the questionnaire for the teachers to correct. The first problem, a fractional equation to be worked, contained two errors. The second problem, a statement problem, was incorrectly set up. The resulting equation, however, was worked correctly. The teachers tended to assign lower marks to a problem incorrectly set up with the basic equation wrong but worked correctly than to a problem with the basic equation already given but worked out incorrectly. Less than half of the teachers assigned no credit at all to the equation worked incorrectly, while more than half of the teachers assigned no credit to the problem with the equation incorrectly set up.

B. A SUMMARY OF THE VALIDITY AND RELIABILITY OF THE TEACHER-MADE TESTS PRESENTED IN CHAPTER IV

Validity: Six points of validity were discussed in Chapter IV in connection with these tests. Perhaps the most important point taken up is the level of difficulty. Each

test was put into one of three categories of degrees of difficulty - difficult, moderate, and easy. Orleans feels that in the selection of test questions, there should be a comparatively small number of exceedingly easy or exceedingly difficult questions and an increasingly larger number of questions that approach average or moderate difficulty.[3] Thus, those tests considered to be of moderate difficulty by the writer are satisfactory on this point according to Table LXII. Tests too easy or too hard are unsatisfactory on this point. Orleans also states that a question is valid statistically when an appreciably larger percentage of the more competent pupils than of the less competent should answer it correctly.[4] For this reason a test too difficult may be more valid than a test too easy.

Nine out of nineteen semester and final tests were considered by the writer to be of moderate difficulty and thus, satisfactory on this point. Of the remaining ten tests, seven were considered to be too easy and three to be too difficult. A teacher seems to be more prone toward constructing a semester or final test that is too easy than that which is too hard. The fear of failing too many pupils may be a factor here. This factor may also be a reason why teachers allow partial credit in their scoring procedure.

---

[3]Jacob S. Orleans, Measurement in Education. (New York: Thomas Nelson and Sons, 1937), p. 48.

[4]Loc. cit.

The picture is reversed somewhat in the case of the six weeks' tests. Thirteen out of seventeen examinations in this category were considered to be of moderate difficulty by the writer. Two examinations were considered to be too easy and two considered too hard. A failure in a six weeks' test does not necessarily fail the pupil for the semester or year. Consequently, the teacher feels safer in setting more difficult examinations, as he does not need to fail a large number of pupils for the semester or year on the basis of these tests. He may set an easier semester or final examination in order to give the failing pupil a chance to "redeem" himself.

The point of second importance is probably the organization of memoriter learnings. Statement or verbal problems call upon the student to organize the algebraic material he has previously learned. The algebra the student has learned in the course will do him little good if he cannot effectively organize the material into practical applications. Seven out of seventeen or less than half of the six weeks' tests were satisfactory on this point. Some teachers tend to shun the problems of application because they feel that these problems are too difficult for the average student to grasp. Nine out of nineteen semester and final tests attempted the organization of memoriter learnings satisfactorily.

A course in algebra should not predominate in too

many "drill" problems to the sacrifice of problems of application. A pupil should be taught how to organize his material early in the course and should be given more problems of that type as he advances. Such problems would not be so difficult as the pupil becomes more acquainted with them. Problems of application teach the pupil to think in an orderly manner, instill more appreciation in the values of the course, and prevent the pupil from becoming too mechanical in his work.

An examination should give a comprehensive measure of the types of items presented. If the problems presented on algebraic factoring, for example, called for varied types of thinking and organization on the part of the student, such problems would be comprehensive. A type of test item may present only one problem and still be comprehensive because that problem requires varied types of thinking. A type of test item may present many problems and not be comprehensive because those problems are too easy or too much of one kind. Comprehensiveness is not related to adequacy of sampling, a point of reliability.

Thirteen out of nineteen semester and final tests each presented a majority of comprehensive types of items. Fifteen out of seventeen six weeks' tests were satisfactory on this point. The semester and final tests presented more types of items than the six weeks' tests. The teacher might give more attention to the types on a six weeks' test than to the greater number of types in a semester or final test.

This is probably why a greater percentage of six weeks' tests are more comprehensive than semester and final tests.

Another important point of validity is social utility. More tests were found to be lacking on this point than on any other point of validity. A test containing local items of interest, practical problems applicable in other fields, and items pertaining to everyday living has social utility. Six out of nineteen semester and final tests were found to be satisfactory on this point. From the seventeen six weeks' tests, only five had some social utility.

Perhaps this is a reason for the student's complaint that he sees no value in algebra. The course is "cut and dried", and most students would only take it because it is required. A student will soon forget what he has learned in a course holding little value as far as he is concerned. More items containing social utility will help to instill in the student more appreciation of the course and give more meaning to the course in general.

Another point of validity, not as important as those already discussed, is the proper limiting of the types measured. A teacher may cram too many types of items into a test and, probably as a result, not measure each type comprehensively or sample each type adequately. It becomes less clear what a test does attempt to measure when too many types of items are included. Most of the tests, fifteen of the semester and final and sixteen of

the six weeks', were found to be satisfactory in this respect.

The least important point of validity taken up in Chapter IV is probably the avoidance of linguistic difficulties. Tests containing many long and difficult words that a teacher should not expect a student to know or tests containing many confusing negative statements are those that would present linguistic difficulties. Tests containing difficult words, such as multiplicand, subtrahend and the like, do not present unnecessary linguistic difficulties as the teacher has a legitimate right to expect students to know these words, providing he has taught their meaning. None of the algebra tests were found to be lacking in this respect.

Reliability: The following table gives a tabulation of the six points of reliability, listed in order of importance, discussed in the preceding chapter.

TABLE LXIV

A RATING OF SIX POINTS OF RELIABILITY
AS SATISFACTORY OR UNSATISFACTORY ON ALL THE TESTS

Points of Reliability	Number of Semester and Final Tests		Number of Six Weeks' Tests	
	Satisfactory	Unsatisfactory	Satisfactory	Unsatisfactory
Objective scoring policy*	9	9	5	11
Usability of test results*	8	10	6	10
Adequacy of sampling	3	16	8	9
Length of test	6	13	6	11
Objectivity	19	-	17	-
Administrability	17	2	16	1

*Two instructors did not make their scoring policy clear. Thus, the total of semester and final tests and the total of six weeks' tests is one short on this point. The usability of test results could not be determined where the scoring policy was not made clear.

C. A SUMMARY OF THE VARIABILITY OF THE PUPIL SCORES

Fourteen of the thirty-six algebra tests discussed in the last chapter also had the pupil scores. The following tables give the standard deviations of the pupil scores and the nature of the instructor's scoring policy on each test.

TABLE LXV

THE STANDARD DEVIATIONS AND INSTRUCTOR SCORING POLICY OF SEMESTER AND FINAL TESTS

Test Number	Standard Deviation	Scoring Policy
27*	15.9	Objective
32	10.2	Not clear
56	34.1	Subjective
62	14.5	Subjective
63	19.9	Subjective
64a	10.3	Objective
	14.2	
69	22.1	Subjective

*Each test in the study was numbered.

TABLE LXVI

THE STANDARD DEVIATIONS AND INSTRUCTOR SCORING POLICY OF SIX WEEKS' TESTS

Test Number	Standard Deviation	Scoring Policy
5	11.0	Objective
12	13.6	Objective
15	19.2	Objective
24	22.6	Subjective
	19.0	
41	20.3	Subjective
	20.7	
44	21.5	Subjective
50	16.5	Subjective

Many of the sets of the pupil scores presented would show variability by the formula

$$V = S.D. \left(1 - r \right)^{\frac{1}{2}}$$

The standard deviations tend to be large and the scoring policy of the instructors tends to be subjective. A subjective scoring policy lessens the basic objectivity of the test and consequently the reliability.

BIBLIOGRAPHY

BIBLIOGRAPHY

A. Books

Greene, Harry A., Albert N. Jorgensen, and J. Raymond Gerberich, Measurement and Evaluation in the Secondary School. New York: Longmans, Green and Company, 1943. 670 pp.

Jordan, A. M., Measurement in Education. New York: McGraw-Hill Book Company, Inc., 1953. 533 pp.

Lee, J. Murray, A Guide to Measurement in Secondary Schools. New York: D. Appleton-Century Company, Inc., 1936. 514 pp.

McCall, William A., Measurement. New York: The Macmillan Company, 1939. 535 pp.

Minnick, J. H., Teaching Mathematics in the Secondary Schools. New York: Prentice-Hall, Inc., 1939. 336 pp.

Monroe, Walter Scott, Measuring the Results of Teaching. Boston: Houghton-Mifflin Company, 1918. 297 pp.

Orleans, Jacob S., Measurement in Education. New York: Thomas Nelson and Sons, 1937. 461 pp.

Trabue, Marion Rex, Measuring Results in Education. New York: American Book Company, 1924. 492 pp.

Wood, B. D., Measurement in Higher Education. New York: Yonkers-on-Hudson: World Book Company, 1923. 337 pp.

B. Periodical Articles

Camp, F. S., "Some 'Marks': an Administrative Problem," School Review, 25: 697-713, December, 1917.

Comin, Robert, "Teachers' Estimates of the Ability of Pupils," School and Society, 3:67, January 8, 1916.

Hughes, W. Hardin, "Analyzing the Ingredients of Teachers' Marks," Nation's Schools, 6:21-25, December, 1930.

Hughes, W. Hardin, "Coordinating Subject Marks and Achievement Marks," Nation's Schools, 10:25-28, October, 1932.

Johnson, F. W., "A Study of High School Grades," School Review, 19:13-24, January, 1911.

Starch, Daniel and Edward C. Elliott, "The Reliability of Grading High School Work in English," School Review, 20:442-457, September, 1912.

Starch, Daniel and Edward C. Elliott, "The Reliability of Grading Work in Mathematics," School Review, 21:254-April, 1913.

Starch, Daniel and Edward C. Elliott, "The Reliability of Grading Work in History," School Review, 21: 676-681, December, 1913.

C. Pamphlets

Montana Educational Directory, State Department of Public Instruction, Helena, Montana, 1952-1953.

APPENDIX

APPENDIX A

EVALUATION IN ELEMENTARY ALGEBRA TESTING

This questionnaire pertains to the tests and measurements program in elementary algebra. Please answer all the questions briefly and return the questionnaire at your earliest convenience in the enclosed self-addressed envelope. Do not sign this questionnaire.

A. General Information

1. What is the enrollment of the high school in which you teach?_____

2. How many sections of elementary algebra do you teach?_____

3. How many students do you have in each section?_____, _____,

_____, _____, _____, _____.

4. How long is the class period for elementary algebra?_____

5. How many years of teaching experience do you have?_____

6. How many years have you taught elementary algebra in the system in which you are now employed?_____

7. Please supply the following information about your basal textbook (or textbooks):

Author	Title	Publisher	Copyright Date

8. Is elementary algebra a graduation requirement as far as your high school is concerned?_____

9. Are any standardized prognostic elementary algebra tests given to the students before they take the course?_____

10. Are any standardized elementary algebra tests other than prognostic given to the students?_____

    a. If so, please name them and state the time of the school year they are given:

    _____

    _____

B. Preparation of your major tests (such as those given at the end of the six weeks or semester)

1. How many major elementary algebra tests do you give your students in the course of a school year?_____

2. Do you generally inform your students before these tests that a test is coming?_____

3. Do you generally select your test items from the basal textbooks used?_____

4. How much time on the average would you say it takes you to make out such a test?_____

C. Preparation of your minor tests (such as daily quizzes)

1. About how many minor tests do you give your students in a typical week?_____

2. Do you generally inform your students before these tests that a test is coming?_____

3. Do you generally select your test items from the basal textbooks used?_____

4. How much time on the average would you say it takes you to make out such a test?_____

D. Administration of your tests

1. How much time do you generally allow your students to take major tests?_____
   Minor tests?_____

2. If you have more than one section, do you give the same major test to each section?_____
   Same minor test to each section?_____

3. Do you usually prepare alternate major tests, so that only every other seated student will be taking the same test?_____

   Do you usually prepare alternate minor tests?____

4. Please indicate by a check mark the form in which you give your major tests:

   On the blackboard_____

   By mimeographed sheets_____

   By dictating the problems_____

   By other methods (name)_____

   Your minor tests:

   On the blackboard_____

   By mimeographed sheets_____

   By dictating the problems_____

   By other methods (name)_____

5. If room permits, do you assign the students alternate seats in the giving of a major test?_____ Minor test?_____

6. Does insuring independent work on the part of your students during a test present a problem?_____

7. How would you suggest such a problem be handled?_____

8. How much emphasis do you place on memorization of material in your testing program?_____

E. Evaluation procedures

1. Do you generally allow the students to correct their own major test papers?_____
   Minor test papers?_____

2. Do you generally allow the students to correct major test papers not their own?_____
   Minor test papers not their own?_____

3. About how much in percentage does each major test count toward a student's final mark for a report period (six or nine weeks)?_____

4. About how much in percentage do all minor tests count toward a student's final mark for a report period?_____

F. The following problems are given for your appraisal. Please grade them any way you wish. The highest possible score is 10 points.

1. The Problem

$$\frac{y+2}{y-3} \div \frac{y+3}{y+2} = \frac{2y^2+7}{y^2-y-6}$$

The Student's Work

Lowest common multiple =

$$y^2 - y - 6 = (y-3)(y+2)$$

$$(y+2)(y-3) \div (y+3)(y-3) = 2y^2 + 7$$

$$y^2 - y - 6 \div y^2 - 9 = 2y^2 + 7$$

$$-y - 6 - 9 = 7$$

$$-15 = 7 - y$$

$$-22 = -y$$

$$y = 22$$

Score you would assign (possible score is 10 points)_____

Please explain why you marked the problem as you did:

2. The Problem

The sum of two numbers is 200. The larger divided by the smaller gives a partial quotient of 3 and a remainder of 8. Find each number.

The Student's Work

Let each number be x, 200 - x
Lowest common multiple = x

$$\frac{200 - x}{x} = \frac{3}{x} \nearrow \frac{8}{x}$$

$$200 - x = 3 \nearrow 8$$

$$200 - 3 - 8 = x$$

$$x = 189$$

$$200 - x = 11$$

Score you would assign (possible score is 10 points)_____

Please explain why you marked the problem as you did:

APPENDIX B

Montana State University
Missoula, Montana
May 6, 1953

High School
, Montana

Dear

I am conducting a survey in the field of tests and measurements as they relate to the marking system used in elementary algebra. Educators today are tending to show more concern toward the study of this field.

Many factors tend to influence the testing program in the school today, such as the textbooks used, the number of sections of students, the number of students in a class or section, the frequency of testing, the preparation and marking of tests, and the weights given to each test.

I have enclosed a questionnaire for your appraisal of various factors as they affect the testing program in elementary algebra. I have attempted to keep the individual items simple in form, most of them requiring only a few words to answer. If any item can not be answered in the space allotted, please do not hesitate to write on the back of the questionnaire.

Your testing program in elementary algebra, no doubt, includes major tests at regular intervals such as six weeks tests and probably minor tests such as daily quizzes. I am making a study of the kinds of tests given by mathematics teachers. To that end I would appreciate it if you would enclose a copy of any major elementary algebra test you have given recently. Please state what section of the book the test covers. Along with this, I would also appreciate the scores on this test made by the pupils in any elementary algebra section or class you may teach. Please designate each pupil only by letter or number.

I am sure that you, as an educator, can see the increasing importance of the field of tests and measurements in the various areas of teaching. You do not need

to give the name of the system in which you are employed. I would appreciate your cooperation in this survey and your return of this questionnaire as soon as your time permits,

If you are interested in the results of this survey, I will be glad to furnish them at your request.

Yours sincerely,

Duane L. Fredricks

Approved: School of Education
Montana State University

APPENDIX C

Montana State University
Missoula, Montana
November 25, 1953

_____ High School

_____, Montana

Dear_____

I am conducting a survey in the field of tests and measurements as they relate to the marking system used in elementary algebra. I am sure that you, as an educator, can see the increasing importance of the field of tests and measurements in the various areas of teaching.

A questionnaire of this type was sent to the elementary algebra instructor at your school last spring. You may not have personally received such a questionnaire at that time, or your schedule of spring activities may not have permitted you to reply.

Various factors tend to influence the testing program in the secondary school, such as the textbook in use, the number of sections of students, the number of students in a class or section, the frequency of testing, the preparation and marking of tests, and the weights given to each test.

A questionnaire has been enclosed for your appraisal of various factors as they affect the testing program in elementary algebra. The individual items are simple in form, most of them requiring only a few words to answer. If any item cannot be answered in the space allotted, please do not hesitate to write on the back of thequestionnaire.

Your program of testing in elementary algebra probably includes major tests at regular intervals, such as six weeks tests, and probably minor tests, such as daily quizzes. I am making a study of the kinds of tests given by mathematics teachers. To that end I would appreciate it if you would enclose a copy of any major elementary

algebra test you have given recently. Would you please state what section of the book the test covers? Also, if possible, I would appreciate the scores made on this test by the pupils in any elementary algebra section or class you may teach. Please designate each pupil only by letter or number.

I would appreciate your co-operation in this survey and your return of this questionnaire as soon as your time permits. If you are interested in the results of this survey, I will be glad to furnish them at your request.

Yours sincerely,

Duane L. Fredricks

Approved: School of Education
Montana State University

APPENDIX D

Montana State University
Missoula, Montana
January 11, 1954

_____ High School

_____, Montana

Dear_____:

A few weeks ago you were sent a questionnaire pertaining to evaluation in elementary algebra instruction and a letter of explanation. This questionnaire was enclosed for your appraisal of the various factors that influence the testing program in elementary algebra. I have attempted to keep the individual items simple in form, many of them requiring but one word to answer.

If it is feasible, I would appreciate it very much if you would enclose a copy of any major elementary algebra test you have given recently together with the scores made on this test by the pupils in any elementary algebra section or class you may teach. Please designate each pupil only by letter or number. Would you please state what section of the book the test covers?

I realize that the semester is drawing to a close and that you may be hard-pressed for time. However, your cooperation would be of great help to me in my survey. If you desire another copy of the questionnaire, I will be glad to send you one together with a stamped, self-addressed envelope at your request. If you are interested in the results of this survey, I will be glad to furnish them.

Yours sincerely,

Duane L. Fredricks
School of Education
Montana State University
Missoula, Montana

Approved: School of Education
Montana State University

APPENDIX E

Montana State University
Missoula, Montana
June 18, 1954

Dear_____:

I have completed the survey in the field of tests and measurements pertaining to the marking systems used in elementary algebra.

A few months ago you were sent a questionnaire to which you replied and requested the results. The questionnaire was divided into three main divisions - the preparation, administration, and evaluation of elementary algebra examinations. The results are presented to you in the form of a summary. The teachers referred to in the summary are the seventy who replied to the questionnaire out of the original 100.

I sincerely apologize if there has been considerable delay in mailing these results to you. A second survey was necessary due to incomplete returns on the first one. This delayed the tabulation of the final results by several months.

Your cooperation in this survey is greatly appreciated. Without your help and that of the other teachers, this study would not have been possible.

Yours sincerely,

Duane L. Fredricks

APPENDIX F

LIST OF MONTANA SECONDARY SCHOOLS COOPERATING IN THE SURVEY

First Class

Anaconda
Billings
Great Falls
Helena

Second Class

Baker
Big Sandy
Browning
Choteau
Columbus
Corvallis
Cut Bank
Ennis
Forsyth
Froid
Fromberg
Glasgow
Hamilton
Harlem
Havre
Hysham
Libby
Malta
Plains
Polson
Roundup
Shelby
Stevensville
Sunburst
Thompson Falls
Three Forks
Victor
Whitehall
Worden

Third Class

Alberton
Antelope
Belfry

Third Class (continued)

Big Fork
Box Elder
Buffalo
Cascade
Circle
Clyde Park
Custer
Dodson
Flaxville
Florence
Frazer
Hingham
Ismay
Joplin
Judith Gap
Lavina
Lima
Manhattan
Melstone
Musselshell
Peerless
Power
Roberts
Roy
Ryegate
Shepherd
Stanford
Sheridan
Superior
Wilsall

County High Schools

Flathead
Missoula
Park
Powell

APPENDIX G

AN ALGEBRA SEMESTER EXAMINATION
CONSIDERED TO BE "EXCELLENT" BY THE WRITER

ALGEBRA
1ST SEMESTER EXAMINATION

NAME_____ DATE_____

Answer

_____1. 1. Given A=2, b=3, c=5, d=1, e=0, and f=6. Find the following:

$$\frac{abc}{df} - \frac{df}{b} + \frac{def}{c}$$

_____2.

2. $16.4 - 3.2.5 \div 15 + 5.3.$

_____3. (a)

_____ (b)

3. Solve for X:

_____4. a. $12X = 72$ b. $1/2\ X = 4$

_____5.

_____6.

4. Solve for X and check:

$$5X + 7X + 6X = 72$$

5. Add:

$$\begin{array}{r} +16 \\ -15 \\ \underline{+12} \end{array}$$

_____7.(a)

_____ (b)

6. Add:

$$-5,6,\ -10,\ 2,4,\ +10,\ -8,\ -3,\ +5,\ -7 =$$

_____8.(a)

_____ (b) 7. Subtract:

a. 841 b. -5.25
-311 ≠4.65

8. Multiply:

a. (-4) (-8) =

b. 3.5 (-2) =

_____9. (a) 9. (a) 24 ÷ (-6) =

(b) (-16) ÷ (-8) =

_____(b)

_____10. 10. Add and check:

5X≠7Y ≠ 3Z ≠ 8 X = 2
2X-9Y -4Z -10 Y = 3
2X≠6Y - Z - 2 Z = 4

11. Simplify:

2a-3a-(2a-b)≠3b-a≠b=

_____11.

_____12.(a)

_____(b)

12. Multiply:

_____13. (a) $4X^2$≠ X ≠ 5 by X ≠ 2

(b) $(4a^2$≠ 6a-7)$(a^2$-a≠ 5)

_____14.

13. Solve for X:
-6 (3-X) = 30

14. Divide:
$2X^2-7XY≠3Y^2$ by X-3Y

_____15.(a)

_____(b)

15. Find the factors:

(a) $4a^2bc≠8abc^2$

(b) 10 a^4≠25a^3 - 45a^2

_____16.(a)

_____(b)

16. Write the products:

_____17.(a) (a) (4H≠5K) (4H-5K)

(b) $(.3a^2-.4b^2)(.3a^2$≠$.4b^2)$

_____(b)

17. Find the factors:

(a) X^2+6X+9 (b) $X^2+2X-15$

_____18. 18. Find the factor: 19. Factor, being sure to obtain prime factors:

$2m^3n - 19m^2n+24mn$

_____19. 20. Same as 19 $32m^4 - 162\ a^4b^4$

$9X^2+42X+49$

_____20. 21. 6% interest on a sum of money is \$720. Find the sum.

_____21.

_____22. 22. If 5 is subtracted from two times a number, the result is 59. Find the number.

_____23. 23. The value of 72 coins is \$5.90. The coins are dimes and nickels. How many are there of each?

_____24.

24. Fourteen years ago John was twice as old as his sister and now the sum of their ages is 40 years. What is the age of each?

_____25.

25. A man walks 9 miles, then travels a certain distance by automobile, and twice as far by train. If the whole trip is 108 miles, how far does he go by automobile? How far by train?

_____26.

_____27.(a)

26. The sum of three numbers is 81. The second number is twice the first, and the third is 6 more than the second. Find the number.

_____(b)

27. (a) At what rate of interest must $8000 be invested to produce an annual income that will pay the rent at $40 a month?

(b) How much must a man invest at 5%, if he already has $12,000 invested at 7%, to give him a total annual income of 6% of his total amount invested?

_____28.(a) Find each man's share of the profits of a business if A receives twice as much as B, and B receives twice as much as C and the profits for the year are $35,000.

_____(b)

_____(c)

(b) The width of a field is to its length as 1 is to 6. If its perimeter is 140 rods, what is the length? (c) The width?

_____29.

_____30.

29. The number of dollars Frank earned in the summer vacation was 30 less than three times the number which Fred earned. Both together earned $130. How much did each earn?

30. A square has the same area as a rectangle whose length is 8 feet more than a side of the square and whose width is 4 feet less. Find the side of the square.

APPENDIX H

AN ALGEBRA SIX WEEKS' TEST

CONSIDERED TO BE "EXCELLENT" BY THE WRITER

ALGEBRA SIX WEEKS TEST JANUARY 13, 1954

Divide:

1. $\frac{12x \div 6y}{2x}$

2. $\frac{8a^2 \div 4a}{2a}$

3. $\frac{24a^2b^2 \div 6ab^2 \div 30a^2b}{3ab}$

4. $\frac{-60m^3n^3 \div 36m^2n^2 - 6mn}{-6mn}$

Long Division:

5. $\frac{n^2 - 7n - 8}{n \div 1}$

6. $\frac{20 - 7x - 6x^2}{4 - 3x}$

7. $\frac{9x^2 \div 25}{3x \div 5}$

Solve following equations:

8. $\frac{b}{9} = \frac{45}{}$

9. $\frac{2w}{3} \div 13 = 7$

10. $5x = 21 - 2x$

11. $3(8x - 2) = 3(4 \div 2x)$

Solve for last letter of alphabet:

12. $y - a = 2b$

13. $ax = -2c$

Solve:

14. $\frac{b}{4} - \frac{b}{8} = \frac{3}{8}$

15. $\frac{x}{7} - \frac{x}{4} = \frac{24}{7}$

16. The length of a playground exceeds twice its width by 25 feet, and 650 feet of fencing are needed to enclose it. Find its dimensions.

17. Bill Jones wanted Sally Smith's telephone number. To tease him, Sally said that ninety added to her age equalled six times her telephone number minus 6060. Bill knew that Sally was 18 years old, but he didn't know enough algebra to call her. Find Sally's telephone number.

18. Let m represent 14. How can you represent the numbers 15, 16, and 17 in terms of m?

19. The sum of two consecutive numbers is 57. Find the numbers.

20. One angle of a triangle is twice as large as another. The third angle contains 5 degrees more than the larger of the two. Find each angle.

21. The length of a certain rectangle is 3 inches greater than the side of a given square. The width of the rectangle is 1 inch less than the side of the square. The area of the rectangle is 31 square inches greater than the area of the square. Find the number of inches in the side of the square.

22. A westbound train leaves the Union Station traveling 55 miles an hour. At the same time, an eastbound local departs at 35 miles an hour. In how many hours will the trains be 270 miles apart?

APPENDIX I

AN ALGEBRA SEMESTER TEST CONSIDERED TO BE "GOOD" BY THE WRITER

ALGEBRA I
FIRST SEMESTER TEST

NAME_____

Add the following:

1) $\begin{array}{r} -4 \\ \underline{5} \end{array}$ 2) $\begin{array}{r} 3 \\ \underline{-3} \end{array}$ 3) $\begin{array}{r} -5 \\ \underline{-7} \end{array}$ 4) $\begin{array}{r} 0 \\ \underline{-5} \end{array}$ 5) $\begin{array}{r} -3 \\ \underline{-7} \end{array}$ 6) $\begin{array}{r} 8 \\ \underline{-9} \end{array}$ 7) $\begin{array}{r} -6 \\ \underline{5} \end{array}$

8) $\begin{array}{r} 4 \\ \underline{-1} \end{array}$ 9) $\begin{array}{r} 9 \\ \underline{8} \end{array}$ 10) $\begin{array}{r} -5 \\ \underline{-6} \end{array}$ 11) $\begin{array}{r} 3x + 2y - 4z \\ \underline{2x - 7 + 6z} \end{array}$ 12) $\begin{array}{r} 3x^2 - 2x + 1 \\ \underline{5x^2 + 4x - 3} \end{array}$

13) $\begin{array}{r} 8a-4b + 3c \\ \underline{9a+5b - 7c} \end{array}$ 14) $\begin{array}{r} 2.6x^2y \\ \underline{8.0x^2y} \end{array}$ 15) $\begin{array}{r} -30y^3 \\ \underline{20y^3} \end{array}$ 16) $\begin{array}{r} 1\frac{1}{4} \\ \underline{-3\frac{1}{2}} \end{array}$ 17) $\begin{array}{r} 1.75e^x \\ \underline{2.25e^x} \end{array}$

18) $\begin{array}{r} -2.73x^5 \\ \underline{-4.27x^5} \end{array}$ 19) $\begin{array}{r} 3a + 3b + 3c \\ \underline{-2a - 2b - 2c} \end{array}$ 20) $\begin{array}{r} 3.2xy \\ \underline{4.6xy} \end{array}$ 21) $\begin{array}{r} 7x + 5y - 7 \\ -2x - 7y + 7 \\ -4x - 2y - 4 \\ \underline{2x - 7y + 2} \end{array}$

22) $\begin{array}{r} 4a + 3b - 5c \\ \underline{2a - 8h - 5c} \end{array}$ 23) $\begin{array}{r} 4x^2 - 2x + 3 \\ -5x^2 + 7x - 9 \\ \underline{x^2 - 6x + 9} \end{array}$ 24) $\begin{array}{r} -2\frac{1}{2}xyz \\ +4\frac{1}{2}xyz \\ -5\frac{1}{4}xyz \\ \underline{+3\ xyz} \end{array}$

25) $\begin{array}{r} 2x + 3y \\ -4x - 7y \\ \underline{-9x + 6y} \end{array}$

Subtract the following:

1) $\begin{array}{r} 3x \\ \underline{-3x} \end{array}$ 2) $\begin{array}{r} 3x \\ \underline{3x} \end{array}$ 3) $\begin{array}{r} -3x \\ \underline{-3x} \end{array}$ 4) $\begin{array}{r} -3x \\ \underline{3x} \end{array}$ 5) $\begin{array}{r} 13a \\ \underline{-4a} \end{array}$ 6) $\begin{array}{r} 15xy \\ \underline{18xy} \end{array}$

7) $\begin{array}{r} 21s \\ \underline{-13s} \end{array}$ 8) $\begin{array}{r} 2.73a^2b^3 \\ \underline{-6.92a^2b^3} \end{array}$ 9) $\begin{array}{r} -3\frac{1}{4} \\ \underline{-1\frac{1}{2}} \end{array}$ 10) $\begin{array}{r} 475x^0 \\ \underline{-525x^0} \end{array}$ 11) $\begin{array}{r} -3215 \\ \underline{+4117} \end{array}$

12) $\begin{array}{r} -819x \\ \underline{-413x} \end{array}$ 13) $\begin{array}{r} 6x - 9y + 10z \\ \underline{4x - 5y - 3z} \end{array}$ 14) $\begin{array}{r} 4x^2 - 6xy + y^2 \\ \underline{3x^2 \qquad -6y^2} \end{array}$

15) $\begin{array}{r} 13x - 14y \\ \underline{- 7x + 18y} \end{array}$ 16) $\begin{array}{r} 8x - 7y + 8z \\ \underline{-9x + 4y - 8z} \end{array}$ 17) $\begin{array}{r} 16a+ 13b - 14c \\ \underline{-17a- 12b + 8c} \end{array}$

18) $a^2 - 2a + 1$ / $2a^2 + 6a - 6$ 19) $14x - 17y$ / $-9x + 3y$ 20) $4a - 4b - 4c$ / $-2a + 2b + 2c$

21) From $9x^2 - 4x - 3$ subtract $3x^2 + 4x - 2$.

22) From $a - b + c$ subtract $2a - 2b + 2c$.

23) Subtract $8x - 7y - 6z$ from $3x - 2y - z$.

24) Subtract $13x - 19y + z$ from $4x + 3y - 6z$.

25) Subtract $8a - 1$ from $8a - 1$.

Multiply the following:

1) -4 / 3 2) 4 / 3 3) -4 / -3 4) 4 / -3 5) $-5z$ / 4 6) $-17a$ / $-3z$

7) 12 / 13 8) -15 / -16 9) $-2\frac{1}{2}$ / $3\frac{1}{4}$ 10) -5 / -3 11) $-1\frac{1}{2}$ / $-3\frac{1}{2}$

12) $3cd_2$ / $-\frac{1}{2}cd^2$ 13) $3z_3$ / $-4z^3$ 14) x^3 / x^6 15) x^{n-1} / x^n 16) $4(4a+3b-5c)$

17) $6(2x^2 - 3x - 6)$ 18) $3(x-3y)$ 19) $5x(a-4)$

20) $-3(2x-3y+4z)$ 21) $7(x^2y-y)$ 22) $x + 2$ by $x + 1$

23) $x^2 + 2x + 3$ / $x + 2$ 24) $x^2 + 2x + 1$ / $x + 1$ 25) $x^2 + xy + y^2$ / $x - y$

Divide the following:

1) -3 by -1 2) -3 by 1 3) 3 by -1 4) 3 by 1

5) 16 by -3 6) -15 by 2 7) $3\frac{1}{2}$ by $\frac{1}{2}$ 8) -10 by 1

9) -10 by -2 10) -34 by -4 11) x^8 by x^4

12) x^4 by x^2 13) $8a^2b$ by $2a$ 14) $-10x^5y$ by $-2x^3y$

15) $9xyz^3$ by $9xz^2$ 16) $10m^2n$ by $5mn$ 17) $12m^4n^4$ by $-2m^3n^3$

18) $.9hk^2$ by $.3hk^2$ 19) E^x by E^y 20) $xy + xz$ by x

21) $B^3 - 3B^2$ by B^2 22) $\frac{1}{2}x^2 + \frac{1}{2}x$ by $\frac{1}{2}x$

23) $x+3\overline{)x^2 + 10x + 21}$ 24) $x-8\overline{)x^2 - 11x + 24}$

25) $a + b)\overline{a^2 + 3ab + 2b^2}$

Solve the following equations:

1) $5x = 10$ 2) $2x = 8$ 3) $3x + 2x + 4x = 45$ 4) $3x = 9$

5) $5x + 7x + 6x = 72$ 6) $\frac{x}{4} = 5$ 7) $\frac{x}{2} = 20$ 8) $90 = \frac{x}{8}$

9) $4 = \frac{x}{7}$ 10) $x + 5 = 10$ 11) $x + 12 = 15$ 12) $x + 25 = 50$

13) $x + 7x = 12\frac{1}{2}$ 14) $x + \frac{1}{2} = 9\frac{1}{2}$ 15) $x-15 = 25$

16) $x-40 = 60$ 17) $x-9 = 50$ 18) $x-12 = 2$ 19) $x-\frac{1}{2} = 10$

20) $x-2.5 = 12.5$ 21) $5x + 4 = 34$ 22) $6x + 9 = 45$

23) $2x + 3 = 19$ 24) $\frac{1}{2}x - 4 = 6$ 25) $2.5x - 6 = 19$

26) $4x - 5 = 23$ 27) $108 = 7x - 28)$ $7x - 12 = 9$

29) $10x = 35$ 30) $\frac{1}{2}x = 15$ 31) $2.5x = 50$ 32) $\frac{x}{3} = 14$

33) $5x + 3x + 4x = 36$ 34) $2x + 2\frac{1}{2}x = 90$ 35) $x + 8.3 = 19.6$

36) $\frac{1}{4}x + 3 = 20$ 37) $\frac{x}{4} = \frac{3}{4} + 3n$ 38) $7x + 2x = 10 + 9$

39) $3x = x - 12$ 40) $5 = 2x-7$ 41) $-32 + 4x = 8$

42) $-9x = 3x - 36$ 43) $20 = -3x + 26$ 44) $-9x = 3x + 36$

45) $5x - 1 = 3x + 7$ 46) $y - 6 = 5y - 18$ 47) $15 - 8 = 3x - 1$

48) $-4 + x = -5x + 32$ 49) $27x + 15 + 12x + 3 - 7x = 15 + 6x - 11 + 12x$

50) $11x + 13x - 4x - 5 = 6x + 7 + 13x - 5$

APPENDIX J

AN ALGEBRA SIX WEEKS' TEST
CONSIDERED TO BE "GOOD" BY THE WRITER

Oct._____ '51

1st. 6 weeks

NAME_____

1. Write the formula for the perimeter of a rectangle._____

2. Indicate the sum of six and three._____

3. Indicate ten less than y._____

4. What are the factors of 14?_____

5. What is the name given to the 4 in the expression $4y^2$?_____

6. What does a·b mean?_____

7. Find the value of 6y if y equals 5._____

8. Write $4y^3$ without using the exponent._____

9. Find the value of $2a^2$ if a equals five._____

10. Is 4a + 5b a monomial, binomial or trinomial?_____

11. Combine like terms in 4a + 10b - a - 2b._____

12. Simplify $8 \div 2 + 6$.

13. Evaluate $x^2 + y^2 - 10$ when x equals 6 and y equals 3._____

14. Find the number of inches in k feet.

15. Solve for x: $7x + 2x = 54$.

16. Solve for y: $7 + 2y = 4$.

17. One number is five times as large as another. Their sum is 3222. Find the numbers. (Show your equation and work in the space just below.)_____

18. What is the cost of seven pencils at y cents each?

19. How much do y sacks of spuds weigh if each sack weighs 90 pounds?

20. A rectangle is five times as long as it is wide, and its perimeter is 246 feet. What are its length and width? Show your equation and work.

21. How many terms are there in the expression: $a - 4b + 6c - 2$?

22. Write 4yyy using an exponent.

23. What is the coefficient of x in the expression $7x^2 - 5$?

24. If x equals 12 and y equals $\frac{1}{2}$, find the value of $x - 10y$.

25. Which are like terms in the expresion $4x^2 - 4x^3 + 10xy - 10x^2$?

26. In the formula $r = 3x - 2$, what does r equal when x equals 4?

27. How much is - 6 plus a - 7?

28. If you add + 8 and a - 3 what is the result?

29. If you add a + 3 and a - 8 what is the sum?

30. Find the sum of the signed numbers $-5 + 4 - 8 + 2$.

31. Simplify: $a + 3x + 6a - 9x$.

32. Find the value of $2ac^2$ if a equals 3 and c equals 4.

33. How would you represent five more than twice a certain number?

APPENDIX K

AN ALGEBRA SEMESTER EXAMINATION CONSIDERED TO BE "FAIR" BY THE WRITER

ALGEBRA SEMESTER

I. Solve for 2 unknowns:

(1) $4a + 3b = -1$, $5a + b = 7$ (2) $r - 6s = -10$, $2r - 7s = -15$ (3) $4x + y = 5$, $2x + 4y = 13$

II. The number of two digits, the tens digit is 7 more than the units digit. The sum of the digits is 11. What is the number?
t = tens digit.
u = units digit.

III. Factor:
(1) $5x^2 - 5y^2$ (2) $ax^2 - 6ax + 9a$ (3) $5r^2 + 10r - 15$

IV. Solve equation by factoring:

(1) $x^2 - 9x + 20 = 0$ (2) $4m^2 + 5m - 6 = 0$

V. Reduce to lowest terms:

(1) $\frac{x}{y} \div \frac{x}{y^2}$ (2) $\frac{2m - 2n}{5} \div \frac{4m - 4n}{10}$

(3) $\frac{x^2 - y^2}{x^2} \div \frac{2x + 2y}{x}$

VI. Solve equation:

(1) $\frac{x}{2} + \frac{2x}{3} = 14$ (2) $\frac{3x}{4} - \frac{x}{2} = 2$

(3) $\frac{y - 8}{5} + 1 = \frac{y - 3}{10}$

VII. Solve for X:

(1) $5x = m$ (2) $ax = b$ (3) $ax - b = 0$

(4) $\frac{x}{3} = a$

VIII. Factor and solve for X:

(1) $x^2 + 5x + 6 = 0$ (2) $y^2 + y - 6 = 0$

(3) $6x^2 = 3 - 7x$

IX. Take the square root:

(1) $\sqrt{73.4}$ (2) $\sqrt{55.0000}$ (3) $\sqrt{69.5800}$

X. Reduce radicals:

(1) $\sqrt{125}$ (2) $\sqrt{99}$ (3) $\sqrt{\frac{5}{8}}$ (4) $\sqrt{\frac{5}{12}}$

(5) $\sqrt{\frac{4}{3}}$

Square root $\sqrt{2} = 1.41$ $\sqrt{3} = 1.73$ $\sqrt{5} = 2.23$ $\sqrt{6} = 2.44$

(6) $8-2\sqrt{5}$ (7) $-11 + 4\sqrt{3}$ (8) $\frac{3}{5} - \frac{1}{5}\sqrt{2}$

(9) $2 - \sqrt{3}$ (10) $5 - \sqrt{6}$

APPENDIX L

AN ALGEBRA SIX WEEKS' QUIZ CONSIDERED TO BE "FAIR" BY THE WRITER

To p. 48-Milne ALGEBRA QUIZ

Name_____

You must solve numbers 3, 10, 12 if able. Choose any other seven to make a total of 10. Show all work neatly on separate paper, with correct set-ups and substitutions, etc. This separate paper must be handed in with this test, so do your scratch work on scratch paper. Write the correct answers on the blanks provided for them on t his paper. Be sure to copy accurately.

1. The earnings of a factory tripled each year for three years. The earnings for the third year alone were 10,800 dollars. How much were the earnings for the FIRST and SECOND years?
   Earnings FIRST year_____SECOND year_____

2. John had 15¢ less than James. James had 25¢ less than Henry. But Henry had twice as much as James. How much had each?
   John had_____; James had_____; and Henry had_____.

3. If A = 0, B = 3; C = 5; D = 4; and E = 6; find the value of:

   $[2 - AB(D^2 - C)] \neq C^2 \neq E \div 2(3 - 2A)^2$ ____Ans.

   Show work here:

4. Add: $5A - 6(B \neq C) \neq 7$, $5(B \neq C) - 6A - 4$, $8A - 9(B \neq C) - 9$, $3A \neq 2 - 5(B \neq C)$. Answer is _____

5. Add: $7(X \neq Y - 1)$, $2B(X \neq Y - 1)$, $-5(X \neq Y - 1)$, $3B(X \neq Y - 1)$. Answer is _____

6. A man earned daily for 8 days four times as much as he paid for his board each day. Besides the 8 days that he worked, he was idle three days. If he had $52.50 left after receiving his wages and paying his board bill, how much did he receive per day, and what was his board per day?
   His daily wages were_____; his daily board was_____.

7. (a) Subtract $14X^2Y^2 - 8XY^3 - 4XY$ from $8X^3Y^2 + 10XY$.
Answer is _____

(b) From $CX - 14ABY + 7A^2B^2$ subtract $9X - 14ABY + 15A^2B^2$.
Answer is _____

8. Simplify by removing parenthesis: (Show work on separate sheet.)
$3A - (2A + 1) + \{A - 3 - (3-\overline{4-A})\}$. Answer is _____

9. Find the value of X when: $7X - 36 - 10X + 12 = 80 - 33X + 46$. X = _____.

10. Three boys had together 85¢. James had 15¢ more than John, and Henry had 25¢ more than James. How much had each?
John's = _____; James = _____; Henry's = _____.

11. (a) Multiply $A^2 + B - 2C$ by $A^2 - B - C$.
Answer is _____

(b) Multiply $X^{(a} - 1) - X^aY^b + Y^{b-1}$ By $X^{1-a} - Y^{1-b}$.
Answer is _____

12. Multiply: $X^{-2c+p+2} - X^{p-2c} Y^{2+p-3c} + Y^{p+4-3c}$ by $X^{2c+p} - Y^{3c-4-p}$. (Can be worked right here - ans. below.)

APPENDIX M

AN ALGEBRA SEMESTER EXAMINATION CONSIDERED TO BE "POOR" BY THE WRITER

SEMESTER EXAMINATION

Do not write on this test. Use your own paper.

1. Tell whether each of the following is a monomial, a binomial, or a trinomial:

   (a) X (b) 5X (c) $7XY \neq Y^2$

2. Write the following using exponents:

   (a) XXX (b) MMMM (c) 6YYY (d) CCCD

   (e) 3AAAA (f) 4hh - 5hhh

3. One number is 5 times as large as another, and their sum is 324. What are the two numbers?

4. A certain number plus 9 times itself equals 80. What is the number?

5. The perimeter of a triangle is 210 in. Find the lengths of its sides if the 3rd side is $1\frac{1}{2}$ times as long as each of the others.

6. Find 4 consecutive integers whose sum is 206.

7. The sum of 4 consecutive odd integers is 256. What are the integers?

Problems 8 thru 12: Add:

8. $5X^2$
   $\underline{5X^2}$

9. $6X^2 \neq 6X \neq 6$
   $-5X^2 \neq X - 1$
   $\underline{X^2 - 2X \neq 1}$

10. $Y \neq 1^0$
    $\underline{-Y - 1}$

11. $8X^3 \neq 2X^2 - 4X - 6X^2 - X^3 - X \neq 3X^2 =$

12. $\frac{Y}{2} \neq \frac{Y}{3} =$

Problems 13 thru 17: Subtract:

13. $\begin{array}{r} 5X^2 \\ \underline{5X^2} \end{array}$ 14. $\begin{array}{r} 5X^2 \\ \underline{-5X^2} \end{array}$ 15. $\begin{array}{r} -5X^2 \\ \underline{5X^2} \end{array}$ 16. $\begin{array}{r} -5X^2 \\ \underline{-5X^2} \end{array}$

17. $\begin{array}{r} X^2 - 5XY + Y \\ \underline{-X^2 - 2XY + Y} \end{array}$

Problems 18 thru 21: Multiply:

18. $\dfrac{5X^3 + 3X^2 - X}{-X^2}$ 19. $\dfrac{M^2N + MN^2 + N^3}{2\ MN}$

20. $X^4 \cdot X^2 =$ 21. $X^8 \cdot X^7 \cdot X^2 =$

Problems 22 thru 27: Divide if possible:

22. $\dfrac{X^5}{X^5} =$ 23. $\dfrac{21C^4}{3C} =$ 24. $\dfrac{4C}{8C} =$ 25. $\dfrac{X + Y}{X} =$

26. $\dfrac{X^2 + XY}{X} =$ 27. $\dfrac{24X - 12Y}{-6} =$

Problems 28 thru 31: Remove parentheses and simplify:

28. $8 + (N + 7) =$ 29. $M + 3(M + N) =$

30. $5X - (X - 5) - 3 =$ 31. $4Y + 2(Y-3) - 2(Y-3) + 1 =$

APPENDIX N

AN ALGEBRA SIX WEEKS' TEST
CONSIDERED TO BE "POOR" BY THE WRITER

ALGEBRA 1

2nd Six Weeks
2nd Semester

This test covered: two unknowns and special products and factoring.

1. $2x - y = 5$
   $6x - 3y = 24$
   are simultaneous, inconsistent, dependent equations.

2. Solve for x and y:

   $5x - 9y = 37$
   $x = y + 5$

Multiply the following:

3. $x^2 (x^3 - ax + b)$

4. $(2n - 3)(n - 6)$

Factor the following into prime factors:

5. $x^2 + 2xy + y^2$

6. $x^2 + 5x + 4$

7. $m^3 + 7m^2 + 12m$

8. $x^4 - 81$

9. $6x^2 - 5x - 6$

10. $3x^2 - 3x$

Solve for the literal number:

11. $(s + 12)(s - 4) = s^2 + 16$

12. $(x + 2)(2x - 1) - (x - 3)(x + 5) = 11 + x^2$

13. A candy maker mixed a grade of candy made to sell at 45¢ a pound with another grade made to sell at 85¢ a pound to make a mixed candy to sell at 67¢ a pound. How many pounds of each grade should he use to make 100 pounds of the mixed grade?

14. The sum of two numbers is 30. If 3 times the larger number is subtracted from four times the smaller, the result is 1. Find the numbers.

Any 10 of the 14 were done.