Indiana State University

Sycamore Scholars

7-1-1946

A comparative study of subjective measurement in general science

Elizabeth A. Ridenour
Indiana State University

Follow this and additional works at: https://scholars.indianastate.edu/etds

A COMPARATIVE STUDY OF SUBJECTIVE MEASUREMENT
IN GENERAL SCIENCE

By

Elizabeth A. Ridenour

Contributions of the Graduate School
Indiana State Teachers College
Number 534

Submitted in Partial Fulfillment
of the Requirements for the
Master of Science Degree
in Education

1946

The thesis of ___ Elizabeth A. Ridenour ___,

Contribution of the Graduate School, Indiana State

Teachers College, Number 534, under the title ___

___

"A Comparative Study of Subjective Measurement

in General Science"

is hereby approved as counting toward the completion

of the Master's degree in the amount of _8_ hours'

credit.

Committee on thesis:

Geraldine Shontz

Inez Morris

Walter O Shriner, Chairman

Representative of English Department:

Edward M. Gifford

Date of Acceptance _July 8, 1946_

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I .

THE PROBLEM AND DEFINITIONS OF TERMS USED

Measurement has always held an important place in the field of education. One of the most essential factors in the teaching process has been that of measuring the results of the teaching. Before the eighteenth century this was done largely by means of the oral examination. Then came the paper-pencil tests which have existed through more than a century and continue to be a necessary part of our modern school procedure.

It wasn't until the second decade of this century that anyone challenged that part of the great tradition of the American public school, - that the teacher's judgement of a student is correct. The criticisms which were advanced were valuable because they led to a new interest in, and a more scientific approach to the important problem of measurement. Chapter II of this study is devoted to these various challenging studies.

I. THE PROBLEM

Statement of the problem. It was the purpose of this study (1) to discover whether or not teachers have a wide variability in the grading of the subjective type of examination paper; (2) to discover whether or not science

teachers tend to be more scientific in their methods of grading than are other teachers; and (3) to show the importance of the comparison factor in grading.

Importance of the study. The several studies that were made during the early part of this twentieth century helped much in arousing an interest in the problem of measurement. The measurement movement thus begun was given great impetus by the Army testing program of World War I.

Walter S. Monroe,[1] educational research director of the University of Illinois, reports that by 1920 educational measurement might have been considered as being in the period of adolescence. Test construction was in the "pioneering" stage with objectivity of scoring a necessary requirement. During the years from 1920 to 1945 educational measurement has developed to the status of early adulthood. Today objectivity is desirable but not necessary from the point of view of validity of the instrument. Now the consideration of validity is relative to the curriculum objectives and the intended function of the test.

One of the important services the measurement movement can render education is in the clarification of its objectives. The Progressive Education Association has

_____
[1] Walter S. Monroe, "Educational Measurement in 1920 and 1945," Journal of Educational Research, January, 1945.

recognized the close relationship between educational objec-
tives and the tools of instruction, one of which is the
test. [2]

The essay examination came into disrepute with the
coming of the new-type objective tests; but with the pres-
ent change in the concepts of measurement, it is again find-
ing an important place in recent literature. Much space is
now given to a discussion of the formulating and the pro-
cedure of marking the essay examination.

In recent literature the terms, "testing" and
"measurement", are being replaced by the new term, "evalu-
ation". There is a marked development in the concept as
to what should be measured. Ability to think, attitudes,
and appitudes are now considered as important as memorized
information. Remmers and Gage [3] in their recent book give
this answer to the question, "Why evaluate?" "To enable
the right pupils to receive the right education from the
right teachers." The new concept of the purpose of eval-
uation is to furnish the necessary data for guidance.
The leaders in the field of measurement in 1945 maintained
that there is need for explicit measurement of all aspects

[2] C.C.Ross, Measurement in Today's Schools.(New York:
Prentice-Hall, Inc., 1941), p.23
[3] H.H.Remmers, and N.L.Gage, Educational Measure-
ment and Evaluation. (New York: Harpers, 1943), p.1

of education growth. Paper-and-pencil tests are being

supplemented by interviews, questionaires, and rating

scales.

There has been much written in favor of the objec-

tive measurement. C.C. Ross, however, affirms:

There is no such thing as a wholly objective measure!
Objectivity is always relative, never absolute. Even
the exact sciences (physics and chemistry) have a sub-
jective element. The subjective element is far greater
in education. It is never possible or desirable to
eliminate human judgement in the intelligent interpre-
tation and application of results. [4]

The writer believes that if it can be shown from

this study that the subjective measurement of teachers is

more reliable than was previously claimed in the early

studies, then a more confident use can be made of essay

tests and other types of subjective evaluation, as needed

in the complete program of educational measurement.

II. DEFINITIONS OF TERMS USED

Comparison factor. The term, comparison factor, is

one that has been used only in the more recent studies on

grading. Though it was not considered in the early studies,

the writer feels that it is the basic factor in grading.

A teacher before grading a test usually assigns a

certain value to each question determined by its importance

in comparison with the other questions on the test. An-

---

[4] Ross, op. cit., p. 95

other procedure followed in grading is that of sorting
which also makes use of the comparison factor. The entire
set of papers are sorted into five groups according to their
comparative merit before the grading of individual papers is
begun. The teacher does not set an arbitrary grade on a
single paper but rather estimates or evaluates it in com-
parison with the other papers in the group. Now consider
two teachers who were given a single paper to grade inde-
pendently. It is to be expected that their evaluations of
this single paper would vary considerably. The early studies
and experiments in grading, which will be reported in Chapter
II, verify this fact. But consider these same two teachers
grading independently an entire set of examination papers.
It would not be unreasonable to expect that these teachers
would agree quite closely on the best paper in the group and
the one that was the poorest and on those that were of mid-
dle value. In other words when the comparison factor is con-
sidered by students of measurement, the reliability of
teachers' grading should be increased.

  Average (best) grade. That the reliability of a cer-
tain teacher's marks might be determined, it was necessary
that the most perfect set of marks possible to obtain be
used. In examining the table of the marks given the papers
by the twenty-five teachers, it will be found that one
paper will be given a high grade by some, a low grade by

others and many grades between the extremes. The grade
that most nearly represents the central tendency of the
entire group might be called the best grade which might
have been given the paper. It was for this reason that a
set of means or averages was selected to represent the
best marks for the entire set of papers. The coefficient
of correlation was computed for each set of teacher's
marks with the corresponding set of arithmetic means or
averages.

III. ORGANIZATION OF REMAINDER OF THE THESIS

Literature studied. The second chapter of this
study concerns a number of previous studies dealing with
the accuracy of teachers' marks and also some of the more
recent writings on the subject of measurement and evaluation.

Material used. Twenty-five unselected papers in
general science were taken from an eighth grade examination
given by the writer. These papers were graded by twenty-
five science teachers from various school systems through-
out the state of Indiana and student-teachers under the
direction of Miss Geraldine Shontz of Indiana State Teachers
College, Terre Haute, Indiana. These teachers' marks fur-
nished the data for the study.

Procedure. The methods used in making the statisti-
cal calculations in this study are given in Chapter IV.

The formulas with explanations and sample calculations will be presented to clarify the method.

Using the twenty-five different marks given to each of the twenty-five papers a mean grade was computed for each paper. This was called the average or best grade, as before mentioned, and was assumed to be the most nearly perfect mark given a paper by any of the twenty-five teachers. The formula used for these computations was the one for calculating the mean of an ungrouped series. The actual marks given each paper were then correlated with these best grades. The Pearson product moment formula and the Spearman formula based on rank differences were used in finding the correlation coefficient.

Results. The results of the calculations explained in the preceeding chapter are presented in Chapter V. Herein are given the computation results in tabular form with their interpretation.

Conclusion. Chapter VI, the last of this study, contains the conclusions reached after this study was completed. These are given in two parts. The specific conclusions resulting from this particular study form the first part. The second part contains general conclusions regarding teachers' marks and examinations.

CHAPTER II

REVIEW OF THE LITERATURE

Our hope of educational improvement lies in evaluation. Educational evaluation involves the passing of judgement on the degree of worth-whileness of some teaching process or learning experience and should not have as its sole purpose the basis for giving a mark. Using the language of the modern artilleryman, educational evaluation is a kind of range-finding in contrast to aimless firing.

Harold B. Dunkel [1] believes that in the immediate future testing devices will be in greater demand and use in helping to solve the problems arising out of war and post-war adjustments. It is, therefore, important that the basic philosophy of measurement be more generally understood.

Evaluation has begun to get away from the memorization of information. The supposed high correlation between measures of information in a certain field and the ability to think in that field are now thrown into doubt. The old oral type of test was wholly for memorized information. This has given way to the pencil-and-paper tests of information in use today but other devices are becoming ever

---

[1] Harold B. Dunkel, "Common Misconception about Testing and Evaluation," School and Society, May 29, 1943.

more important. Mental ability is being broken down into various aptitudes.

C. C. Ross [2] gives three stages in the use of tests. (1) Curiosity stage (2) stage of confidence or over confidence and (3) the present stage, one of critical caution. Tests are a means not an end; a tool to be skillfully and intelligently used. Measurement may now be considered to be coming of age. The publication of the first volume in 1938 of the Mental Measurement Yearbooks by Rutgers University Press marked an important milestone in the history of educational measurement.

"The measurement movement in education always has been criticized and it is safe to prophesy it always will be. This is as it should be. Measurement is becoming ever more important for it has spread upward to adults and downward to the pre-school child." [3]

The so-called traditional test or essay examination has been in existance for hundreds of years but the amount of research devoted to it is less than that on the new-type or objective test. The research has been of the negative kind designed to show how poor rather than to show improve-

---

[2] Ross, op. cit. pp. 67-68

[3] S. A. Courtis, "Current Criticisms of Educational Measurement," Review of Educational Research, 8:545-50, December 1938

ment. Most of this evidence has been in the studies of the extreme variability of teachers' marks. There has been only one study, by Meyer, made on a comparison of the old and new type tests. [4]

The Starch-Elliott Study in English. [5] This study was made to determine the range of variation and the reliability of the marks assigned by teachers to examination papers. A single paper was to be graded independently by a great number of teachers. Two English examination papers written by two first-year students in the largest Wisconsin high school were used. Several hundred copies of these two papers were printed and sent to each of two hundred high schools in the North Central Association of High Schools and Colleges to be graded by their principal English teacher. Some were also graded by English students in the University and others by students in a class on Educational Measurement in the University of Chicago. The ranges and medians of these groups were almost identical whether the group doing the grading was a small or large group. The most startling fact of this investigation was the wide range of variation in marks (from thirty-five to forty points) for the same paper.

———————

[4] Remmers and Gage, op. cit., p. 131

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

The poorer paper was given five points above passing by one teacher but below passing by twenty-two of the grading group. Such a variation would have an important bearing upon the question of promotion and retardation if the examination were used as a sole basis. Graphs picturing the grades showed that the same teacher did not grade correspondingly high or low on both papers.

Starch-Elliott Study in mathematics. [6] It had been urged that the marks given on language work would naturally vary considerably because of such contributing factors as spelling, sentence construction and vocabulary. This would be different in marking mathematics which is an exact science. The investigation made by Starch and Elliott a year later in mathematics was a sequel to the one made by them in English. There was only one paper, a geometry examination, used in this study. The principal teacher of mathematics in the high schools participating did the marking. One hundred forty papers were graded and returned. The schools represented in this study had arbitrary passing grades of 70, 75, or 80 rather than a passing grade determined by the type of examination given or class taking it.

---

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

This investigation failed to show any more reliability or lack of variation but rather a more extremely wide variation than the English study did. The probable error was 7.5 in mathematics while in English the probable error of the two papers was 4.0 and 4.8.

It was suggested that the wide variation in marks might be due to the form or make-up and neatness of the papers. Some teachers supposedly disregarded these items while others did not. Another explanation given was that schools doing the grading were so widely scattered and therefore had different standards of attainment. This did not seem to be the real cause of the variation because it was found that four teachers in the same high school assigned grades of 76, 75, 67, and 61 to the paper. It was noted that marks of a single answer to one question varied as widely as those of an entire paper.

Another study was conducted by the same authors in history. Again only one paper was graded by 122 different teachers. Their findings in this study corroborated with those of mathematics and English. It was concluded that variability of marks is not a function of the subject but of the examiner.

The Shriner Study.[7] This study was made in two high

_____

[7] Walter O. Shriner, "The Comparison Factor in the Evaluation of Examination Papers", The Teachers College Journal, 1: 65-74, January 1930.

school departments, sophomore algebra and freshman English.
Twenty-five final examination papers were secured from
these two departments in the Ann Arbor, Michigan High
School. These papers were made under as normal conditions
as possible in classes composed of unselected pupils.

The sets of original papers in all cases were sent
either to teachers of various lengths of experience or to
advanced college students of those subjects. No instruc-
tions as to the method of scoring were given except that the
grades were to be given in per cents based on the scale of
100. The persons grading the set of papers were asked to
designate the grade they would consider a passing mark on
that set of papers. No marks were to be made on the origin-
al set of papers that might influence another grader.

The coefficient of correlation was computed for each
teacher's marks with the corresponding arithmetic mean in
each test. The following conclusions were found:

1. The reliability of the various teachers' marks
   for both subjects was very high. There were only
   three with a correlation less than 0.90 in alge-
   bra and all over 0.85 while in English there were
   none below 0.80 and only six less than 0.90. The
   median correlation coefficient for algebra was
   .946 and .917 for the English with comparatively
   small probable errors.

2. The length of teaching experience was not impor-
   tant in the matter of grading.

3. The greatest difference and fault lay in choosing
   the suitable arbitrary mark as a passing grade.

The Stalnaker Study. [8] Earlier studies had proved
that the ordinary essay-type examinations were not reliably'
read, and recommended the use of objective tests to increase
the reliability of scoring. The purpose of this study made
by the consultant examiner of the College Entrance Examin-
ation Board is to show that essay examinations can be re-
liably and consistently read. "The accurate scoring of
objective tests requires more care than is ordinarily ac-
corded it."

The papers from the examination given each year by
the College Entrance Examination Board in various centers
are sent at once to New York. Here readers assemble from
schools and colleges to evaluate them. About 100 readers
read the English paper in 1937.

The student's name is concealed. A score sheet, pro-
viding space for each part score made on the examination, is
stapled on each answer book. No marks are made on the book-
let itself. After all score entries are made the book is
turned over to a comptometer operator who arrives at the
total score which is converted to an appropriate scale and
reported as the obtained grade.

The grade sheets are removed from a sampling of books,
new ones are attached, and the books are sent back through

---

[8] John M. Stalnaker, "Essay Examinations Reliably
Read", School and Society, 46: 671-672, November 1937.

the reading process. This second, completely independent reading of the book provides a means of checking the reliability of the grading.

The table showed that all the subjects except English (.84) were read with a reliability of over .90 and several, mathematics and chemistry, of .98 and .99 respectively.

Stalnaker concluded that cooperation between technician and reader will result in improved procedures which should raise these reliabilities to even higher figures. Such studies should dispel the idea that essay examination cannot be consistently and reliably read.

The Ashburn Study. [9] This experiment was suggested by a discussion in a meeting of lecturers at West Virginia University. While one was attempting to make all the tests in his course of an objective type, the others suggested that the essay question should be used because of their belief that it tests certain abilities that cannot be measured by other types of questions. Discussion questions on Dante were given to different groups of students and three members of the faculty, authorities on Dante, were asked to read the papers and assign to each a letter grade.

---

[9] Robert Ashburn, "An Experiment in Essay-Type Questions", Journal of Experimental Education, 7: 1-3, September 1938.

At first glance there seemed to be a high degree of agree-
ment. The average grades for the sixty-five papers ran
about the same but the failing grades were not assigned to
the same papers. Correlations were run on them by pairs
with the following results: Professors R and S, .65 ± .05;
R and T, .67 ± .05; and S and T, .69 ± .04.

Ashburn wished to continue the experiment under more
fair conditions. All the students were given the same essay
questions, one on history and the other on English litera-
ture. These had been prepared by committees of three faculty
members each. The correlations of them by pairs were:

	X and Y	X and Z	Y and Z
history	.72 ± .04	.71 ± .04	.61 ± .05
English	.77 ± .03	.81 ± .03	.84 ± .02

The question was raised as to whether there was a
consistent standard used by each member of the faculty in
grading essay tests. Members of the English committee a-
greed to read the papers again after a period of two weeks.
X felt that the papers had deteriorated in two weeks for on
rereading his average grade dropped, Y was in better humor
during the second reading and the grades improved, five
passing who had failed on the first reading. Z considered
the papers worse and found ten more failures on careful
reading. All three used greater care in their second read-
ing but their correlations with each other dropped. The
following conclusions were reached:

1. The passing or failing (the difference between credit and no credit) of about 40 per cent depends not on what the student knows or does not know, but upon _who_ reads the paper.

2. The passing or failing (the difference between credit or no credit) of about 10 per cent depends not on what the student knows or does not know, but upon _when_ the papers are read.

The correlations reported by Ashburn were made on pairs rather than on single grades and thus do not give as true a picture as possible. An average or best set as described in Chapter I could have given truer correlations.

The Lawson Study.[10] Douglas E. Lawson believed that a teacher's marks may have a large effect on a pupil's self confidence. "The teacher who deals in grades may be handling material as lethal as anything sold by the pharmacist or administered by the practitioner-- lethal, that is, if wrongly used."

While teaching an extension class in educational tests and measurements in a German community of southern Illinois, he carried out a study of the grading done by teachers having uniformity of background and professional preparation. Certainly here was a condition in which there was ideal uniformity of the graders. Of the forty-one experienced teachers in his class participating, thirty-five claimed German ancestry; twenty-nine spoke German; thirty-

[10] Douglas E. Lawson, "Teachers' Marks--Tragic and Absurd", _The Educational Forum_, 4:175-179, January 1940.

eight were born in the same county; thirty-four attended
the same elementary and high school, and twenty-two had
their training in the same college. All were teaching with-
in eighteen miles of the county seat and all but one were
teaching in the elementary schools.

Some of the papers to be graded were collected from
actual students in different grades. Others were prepared
with errors deliberately made in them. Mimeographed copies
were made of the nine papers finally selected. These were
to be considered as having been written in a final exam-
ination by pupils in the grades indicated. The teachers
doing the grading did not know it was an experiment. Since
it was one of the assignments in their extension course,
their best efforts would be put forth in the task of grad-
ing.

On an eighth grade essay on American literature the
grades ranged from 100 to 40. Grades from 90 to 10 were
given a second grader's letter. Ranges from 95 to 20 and
100 to 10 were reported on sixth and seventh grade dis-
cussion questions.

The experimenter feels that the teacher, in allow-
ing so much subjective judgement to enter into her grad-
ing, builds a false sense of superiority in which her
school becomes a little kingdom. He advises a greater use
of objective grading and an increased use of standard tests.

Although this study was made much more recently than the Starch-Elliott studies and was intended to be an improvement by using a homogeneous group of graders, little is added to the 1912 and 1913 studies. The method of grading only one paper at a time was still used in this study and this is not a normal method of grading.

The **Overholser Study**. [11] This study used data from the grading of twenty-five unselected papers in plane geometry. Marks were given these papers by twenty-five teachers from various school systems. No instruction as to scoring was given except that per cent was to be used in grading based on the scale of 100. The teachers doing the grading were asked to designate the grade they would consider a passing mark.

The coefficient of correlation was computed for each teacher's marks by comparing it with the average or best paper of the group.

The following facts were revealed by this study:

1. The reliability of the teachers' marks was high. The median correlation coefficient based on grades was 0.917 and 0.903 based on ranks.

2. The frequent listing of the arbitrary 70 or 75 as the passing grade was not so apparant as in previous studies. The passing grades varied as much as 40 per cent.

---

[11] C. Daniel Overholser, "The Significnace of the Comparison Factor in Measurement", Thesis, Indiana State Teachers College, 1940

3. The range of marks assigned to a single paper
was found to be about as great in one case as
any found by Starch and Elliott. The range of
grades is greater in the case of the poorer
papers, the better grades being more compact.

The Brown Study. [12] This was an experiment conduct-

ed with home economics seniors completing the course in

supervised teaching and with graduate students who were

experienced teachers. Seven essay-type examinations were

graded by twelve home economics teachers. A per cent

grade was given based on 100. Four of the seven papers

had a range of 50 or more while the poorest one in the

group had a range of 15 to 45.

This study was reported by Remmers to show that

home economic's papers are as difficult to score accurate-

ly as tests in any other subject field.

---

[12] Brown, Clara M., Evaluation and Investigation in
Home Economics, (New York: F.S. Crofts and Company, 1941)
cited by Remmers and Gage, op. cit. p. 129

CHAPTER III

MATERIAL USED

The examination, a copy of which will be found in Appendix A of this study, was the material used to secure data for this study. The examination was one given by the writer to her general science classes as a test over the first semester's work, and therefore was quite general in nature. There are questions touching upon the various sciences: biology, physics, chemistry, and astronomy.

The examination is largely of the essay-tpye, two questions calling for an explanation and two for a discussion. The greater part of this examination, seven out of the eight questions, are subjective. While in question three a definite scientific principle must be stated, the student is given his own choice in naming an everyday example for this principle. He is also allowed to choose the heavenly body which he will discuss in question six. The method of grading the major part of this examination, because of its very nature, must be subjective. Some teachers will give credit for the type of examples chosen and thoroughness of discussion and explanation, while others will give more credit to a well-worded, concise discussion in which only the pertinent facts are stated.

The fourth and eighth questions of the test are the

objective type. The symbols and formula for the chemical

element and compound must of course be considered object-

ively. But this fourth question has an element of subjec-

tivity because the student is allowed to choose the element

and compound he gives and the chemical change he names.

Teachers in grading this question might also give different

weight to the various parts of the question depending upon

whether they thought the knowledge of symbols of equal or

less importance with the knowing of the difference be-

tween elements and compounds. Question eight is one of

matching, one type of the objective or new-type examination

question.

The examination was hectagraphed and each pupil

was given an individual copy of the test. Since this was

given as an eighth-grade semester examination in the regu-

lar class period, the conditions of the investigation were

entirely normal. The same test was given to all five

sections of the eighth grade. Because mental ability had

been considered in the grouping of the students into

sections, the papers used in this study were not taken

entirely from one section. The twenty-five papers used

were those of unselected pupils, chosen from the several

sections in order that a more normal range of ability might

be had.

The examination answers were written in pencil on

the students' own notebook paper as was their custom in
taking their science examinations. The twenty-five
students whose papers were used in the study were asked
to make two exact copies of their examination papers, writ-
ing in pencil and on only one side of the page. The writer
was desirous that the papers should be a typical set of ex-
amination papers in which the students' own hand writing,
composition, spelling, and arrangement were evident. By
having two complete sets of papers it was possible to have
more than one teacher working on them at the same time and
also to have an extra set in case one became lost in the
handling by so many teachers. Since the questions were
written on loose pages, the teacher doing the grading was
given the opportunity of completing one question on each of
the twenty-five papers before proceeding to the next ques-
tion. This procedure is recommended as one of the steps
in securing greater reliability in the grading of essay-
type examinations.

These papers, prepared as described above, were sent
out consecutively to science teachers throughout the state
of Indiana, who had expressed their willingness to partici-
pate in this study. A copy of the letter which was sent
to science teachers explaining the study appears in Appen-
dix B. The teachers participating in this study were of
varied lengths of experience and degrees of skill. Some

were teachers of elementary science while others taught in
the special fields of science in high school and college.
Several were students in a college science methods course.
No instruction was given as to how the different elements
were to be scored. A letter of general instruction was
sent along with the set of papers to be graded. In this
letter all teachers were cautioned to make no corrections
on the original manuscripts so that subsequent readers
might not be influenced in their marking.

All teachers were requested in submitting their reports
not only to give the marks in terms of per cent based on
the scale of one hundred, but also to rank the students.
The graders were also asked to designate what they consid-
ered the lowest passing mark on the test. The reports of
all this desired information made by all the teachers
are included in Appendix C.

The data used for this study include the independent
marks of twenty-five teachers of science on twenty-five
papers. This represents a total of 625 individual grades
on the reading of papers. Such a number is deemed quite
sufficient for a statistical study and the drawing of
general conclusions.

It might be thought that each teacher knowing he
was participating in a study of grading would give more
than usual care to the evaluation of the papers. In this

case, the results would not be typical of the subjective
grading of teachers. The possibility of this factor's hav-'
ing a determining influence is partially offset by two other
facts. A paper is more easily graded if the corrections
can be indicated on the paper itself. Furthermore, the
teacher under regular classroom conditions, occasionally
makes careless mistakes in his grading. These are detected
and the grade is made more accurate when the test is return-
ed and discussed with the pupils. These two possibilities
may be considered as overbalancing the unusual care of the
participating teachers and bringing the conditions of this
study more closely to the normal grading situation.

CHAPTER IV

PROCEDURE

After the data had been gathered as described in the previous chapter, it was subjected to a statistical study. The information of the marking as reported by the twenty-five teachers in Appendix C was compiled for the first part of the study in Table I, page 27.

In the statistical study the coefficient of correlation was computed for each teacher's set of marks with the corresponding set of arithmetic means or averages. In order to determine the reliability of a given teacher's marks, it was necessary that the most perfect set of marks obtainable be used in that measurement. It was found in each group of twenty-five papers that the arithmetic mean and the median for the group were approximately the same. Because the mean is better for further statistical computation, the set of means or averages was chosen as representing the best marks for the various papers.

The assumption that the average scores represent the best set of marks obtainable for each set of papers is based on the mathematical theory of dispersion. Dr. Shriner[1] computed the coefficient of correlation by correlating the

[1] Walter O. Shriner, "The Reliability of Teachers' Marks," The Mathematics Teacher, 17:426-443, November, 1924

TABLE I

SUMMARY OF THE MARKS GIVEN A SET OF TWENTY-FIVE

GENERAL SCIENCE PAPERS BY TWENTY-FIVE TEACHERS

Paper	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	(Best) Grade
A	94	98	91	95	98	94	85	96	98.5	92	92	98	90	100	93	97	95	94	98	93	90	82	100	100	95	94.3
B	89	95	87	87	88	89	70	98	88.5	92	91	86	82	88	74	96	85	88	92	89	94	95	94	96	90	88.9
C	95	98	94	93	97	90	91	94	89	92	87	95	89.5	99	86	87	95	94	94	83	90	85	96	96	95	92.2
D	90.5	94	94	89	95	88	81	96	95	91	92	87.5	86.5	98	90	92	91	91	92	97	92	93	90	95	98	91.9
E	87	90	79	81	78	80	82	85	88.5	86	89	87.5	84	92	79	90	88	88	85	94	86	91	84	91	90	86.2
F	83	92	77	77	83	77	94	87	81	82	88	88	76	89	66	93	81	94	81	91	75	70	84	95	98	84.1
G	93.5	89	88	94	98	91	95	87	95.5	97	95	100	100	97	70	90	97	94	93	91	90	90	98	96	85	92.6
H	74	88	81	80	77	75	86	81	77.5	77	82	93.5	71	48	60	85	90	82	93	83	75	65	81	88	82	79.0
I	68	80	77	79	76	77	69	81	75	66	75	76	68	82	70	90	83	88	90	78	73	85	79	81	90	78.2
J	78.5	85	71	76	74	79	85	74	69.5	77	84	87.5	68	74	79	80	85	79	82	76	74	82	85	80	85	78.8
K	78	84	81	83	87	84	82	81	72.5	84	86	96	66.5	90	61	91	93	81	79	80	83	90	89	93	87	83.3
L	51.5	79	37	63	44	60	47	75	55	44	65	65.5	37	62	33	75	53	52	74	50	49	49	54	62	75	56.4
M	52.5	65	66	65	70	62	67	71	69.5	75	74	80	51.5	76	60	80	81	78	75	76	72	61	69	77	75	69.9
N	52.5	72	66	69	63	81	56	58	63	65	72	77	46.5	69	56	73	78	70	78	55	71	76	75	78	72	67.7
O	89	76	66	80	79	83	83	76	85	79	92	90	68.5	88	63	79	98	88	92	89	76	73	88	92	80	82.1
P	51	62	48	66	71	62	60	65	58	53	60	63	49.5	66	36	70	72	64	68	65	54	48	60	67	78	60.7
Q	48	60	47	61	52	55	50	52	53	47	80	40.5	39.5	54	29	78	63	55	70	52	50	44	55	65	70	54.8
R	56.5	82	51	60	59	62	60	61	61	64	82	64.5	55.5	74	49	73	78	66	76	65	50	44	61	73	78	64.2
S	47.5	75	57	66	71	60	61	69	69	63	83	68.5	59	69	54	83	90	64	70	62	67	79	58	70	85	68.0
T	64	81	54	68	69	62	51	55	51	60	83	65.5	59	64	58	75	74	59	74	60	64	44	67	64	90	64.6
U	72.5	91	74	75	81	86	81	84	74	78	97	87.5	70	89	71	92	88	80	94	76	80	85	77	86	95	82.6
V	75.5	74	72	80	79	72	71	84	74	78	88	83	85.5	80	81	88	95	84	91	80	86	78	75	82	98	81.4
W	72.5	79	73	82	85	79	89	81	79.5	90	92	93	84	84	69	90	95	74	88	73	76	78	73	81	88	81.9
X	87	74	82	87	98	85	87	83	85.5	91	91	92	88.5	91	74	89	97	92	88	91	85	90	89	100	100	88.7
Y	65.5	55	64	63	79	76	65	71	68	63	83	94	61	76	45	80	91	78	83	89	80	76	55	88	70	72.7

various pairs of teacher's marks. Since there were imper-
fections in both sets of marks being used, this method re-
sulted in lower correlations.

The means were calculated by use of the formula

$$M = \frac{\Sigma X}{N} .$$

This is the method for computing the mean for ungrouped
series. For example, for paper A all the marks assigned to
it were added so that the total gave the term ΣX. N is the
number of papers, twenty-five.

A derived form of the Pearson product moment formula
was used in making the computations of the coefficients of
correlations. This derived form for use with a calculating
machine is:

$$r = \frac{N \cdot \Sigma xy - \Sigma x \cdot \Sigma y}{\sqrt{[N\Sigma x^2 - (\Sigma x)^2][N\Sigma y^2 - (\Sigma y)^2]}}$$

The term r_{xy} is the coefficient of correlation between two
items x and y. In this study x represents the grades as-
signed to each paper by one teacher, and y represents those
found by computation to be the mean or average (best) grades.
As shown in Table I, page 27, the x represents the marks in
the columns numbered from one to twenty-five, and y repre-
sents the column of average (best) grades.

Suppose the calculation of the correlation of Teacher 10
be used as a sample of the correlation procedure. The number
of examination papers, N, is 25. The sum of the x's, Σx, is
1886.00. The sum of the y's, Σy, is 1945.20. All the x's

were squared individually, and their sum, x^2, was

147,700.00. All the y's were squared individually, and

their sum, Σy^2, was 154,631.04. Each x was multiplied

by its corresponding y value, and the sum of all of these

products, Σxy, was 150,786.90. The square of the sum of

all the x's, $(\Sigma x)^2$, was 3,556,996.00. The square of the

sum of all of the y's, $(\Sigma y)^2$, was 3,783,803.04. Substitut-

ing these values in the equation on page 28, we have

$$r_{xy} = \frac{25 \times 150,786.90 - 1886 \times 1945.20}{\sqrt{(25 \times 147,700. - 3,556,996)(25 \times 154,631.04 - 3,783,803.04}}$$

and this gives the value to r_{xy}, the coefficient of corre-

lation, of 0.958.

This same method was used to calculate the coefficient

of correlation for each of the twenty-five teachers.

It is always desirable to know the probable errors of

each correlation coefficient. For these calculations the

following formula was used:

$$P. E._r = \frac{.6745(1-r^2)}{\sqrt{N}} .$$

A sample calculation is not given for this because the

operation is comparatively simple. The r represents the

correlation coefficient and N represents the number of

cases used, which was twenty-five.

As a further means of comparison, the correlation

coefficients for the twenty-five teachers were found using

the ranks of the papers. From the original report sheets

submitted by the individual teachers and to be found in
Appendix C of this study, the information for this pur-
pose was compiled as shown in Table II, page 31.

The formula used for these computations was Spearman's
formula based on rank differences,

$$p = 1 - \frac{6 \sum (V_X - V_Y)^2}{N(N^2 - 1)},$$

where V_X and V_Y are the ranks of the x and y terms, respec-
tively, and N is the number of cases, twenty-five.

The V_Y term, shown as M in the last columns of Table II,
page 31, was found as the mean was found in Table I, page
27; in this instance, however, the mean was in each case
reduced to the nearest whole number so that the set of
averages or means was a list of whole numbers.

All computations have been checked for accuracy.

TABLE II

SUMMARY OF THE RANKS GIVEN A SET OF TWENTY-FIVE

GENERAL SCIENCE PAPERS BY TWENTY-FIVE TEACHERS

Teacher

Paper	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	M
A	2	1	3	1	1	1	7	2.5	1	3	3	2	2	1	1	1	3	1	1	3	3	6	1	1.5	6	2
B	5	3	5	5	4	4	12	1	5.5	3	4	2	8	9	7	2	9	4	4	5	1	1	4	4	9.5	5
C	1	2	1	3	2	3	3	4	4	3	2	3	3	2	3	13	3	1	2	6	3	5	3	4	6	3
D	4	4	2	4	3	5	10	2.5	3	5.5	3	8	5	3	2	4	6	3	4	1	2	2	5	6.5	3	4
E	6	7	9	9	10	9	9	7	5.5	8	5	8	7	5	5	7	8	4	8	2	4	3	9	10	9.5	7
F	7	5	10	14	7	14.5	2	5.5	9	10	6	7	9	8	13	3	11	1	11	4	9	11	9	6.5	3	8
G	3	8	4	2	1	2	1	5.5	2	1	2	1	1	4	10	10	2	1	3	4	3	4	2	4	15	4
H	11	9	8	10	11	17	6	12.5	11	14.5	11	5	10	20	16	14	7	6	3	6	9	12	10	11.5	17	11
I	13	14	11	13	12	14.5	13	12.5	12	17	13	13	13	11	10	8	10	4	6	7	11	5	11	15.5	9.5	11
J	8	10	15	15	13	12.5	7	7	16.5	14.5	9	8	13	14	5	16	9	9	10	8	10	6	8	17	15	11
K	9	11	7	7	5	8	9	12.5	15	9	8	3	14	7	15	6	4	7	12	7	6	4	6	8	13	9
L	18	17	25	22	20	23.5	21	16	23	25	16	15	22	18	24	22	16	17	16	15	18	14	21	25	21.5	20
M	16	22	16	21	15	20.2	14	18.5	16.5	16	14	11	18	13	16	17	11	10	15	8	12	13	15	19	21.5	16
N	16	21	17	17	17	11	18	23	20	18	15	12	20	15	19	23	12	12	13	13	13	9	13	18	23	16
O	5	16	18	12	9	10	8	15	8	11	3	6	12	9	14	19	1	4	4	5	8	10	7	9	18	10
P	19	23	23	19	14	20.2	17	21	22	23	17	17	19	16	23	25	14	14	18	10	16	15	18	22	19.5	19
Q	20	24	24	24	19	25	20	25	24	24	12	18	21	19	25	20	15	16	17	14	17	16	20	23	24.5	20
R	17	12	22	25	18	20.2	17	22	21	19	11	14	17	14	21	24	12	13	14	10	17	7	17	20	19.5	17
S	21	18	20	20	14	23.5	16	20	18	20.5	10	16	16	15	20	15	7	14	17	11	14	16	19	21	15	16
T	15	13	21	18	16	20.2	19	24	25	22	10	15	16	17	18	21	13	15	16	12	15	16	16	24	9.5	17
U	12	6	12	16	8	6	10	8.5	13.5	12.5	1	8	11	8	9	5	8	8	2	8	7	5	12	13	6	9
V	10	19	14	11	9	18	11	8.5	13.5	12.5	6	10	6	12	4	12	3	5	5	7	4	8	13	14	3	10
W	12	15	13	8	6	12.5	4	13.5	10	7	3	5	7	10	12	9	3	11	7	9	8	8	14	15.5	12	9
X	6	20	6	6	1	7	5	10	7	5.5	4	6	4	6	7	11	2	2	7	4	5	4	6	1.5	1	6
Y	14	25	19	23	9	16	15	18.5	19	20.5	10	4	15	13	22	18	5	10	9	5	7	9	20	11.5	24.5	14

CHAPTER V

THE RESULTS

The results of the various gradings made by the twenty-five teachers on the set of eighth grade examination in general science, compiled from the individual teacher reports found in Appendix C, are to be found in terms of their per cent marks in Table I, page 27, with the computed means or averages for each paper.

The first column of Table I, page 27, gives the pupil papers according to an alphabetical arrangement from A to Y, inclusive. Columns one to twenty-five, inclusive, give the per cent grades assigned to the corresponding papers by the various teachers who are designated by the numbers one to twenty-five. The last column gives the arithmetic means for the papers. As explained previously, these sets of mean grades were used as the best marks obtainable for the papers.

A graph based on the grades as found in Table I, page 27, shows the distribution of grades given by the twenty-five general science teachers. In Graph I, page 33, there are twenty-five silhouettes, one for each teacher respectively. In comparing the grades given the set of papers by the various teachers, the per cent grades are grouped in five divisions similar to the five letter grades commonly

used in grading. The first bar pictures the per cents
100 - 95. The other bars in order represent per cents of
94 - 88, 87 - 77, 76 - 65 and below 65.

The results of the various gradings made by the twenty-
five teachers on the papers, compiled from the individual
teacher reports found in Appendix C, are to be found in
terms of their ranking marks in Table II, page 31, with
the computed means or averages for each paper.

The first column of Table II, page 31, gives the pupil
papers according to an alphabetical arrangement from A to
Y, inclusive. Columns one to twenty-five, inclusive, give
the ranks assigned to the corresponding papers by the
various teachers who are designated by the numbers one to
twenty-five. The last column gives the arithmetic means for
the papers. These sets of mean ranks were used as the best
ranks obtainable for the papers and were reduced to the
nearest whole numbers.

Table III, page 35, presents summaries of important
relationships based on the per cent marks as given in Table I,
page 27, and on the ranks as given in Table II, page 31. The
first column of Table III, page 35, designates the teacher;
the second column gives the coefficient of correlation be-
tween the teacher's set of marks and the best set; the third
column gives the probable error of the correlation coefficient;
the fourth column the lowest passing grade as indicated by

SUMMARY OF THE RELATIONSHIPS BASED ON THE MARKS

GIVEN TWENTY-FIVE GENERAL SCIENCE EXAMINATIONS PAPERS

CLASSIFIED ACCORDING TO TEACHERS

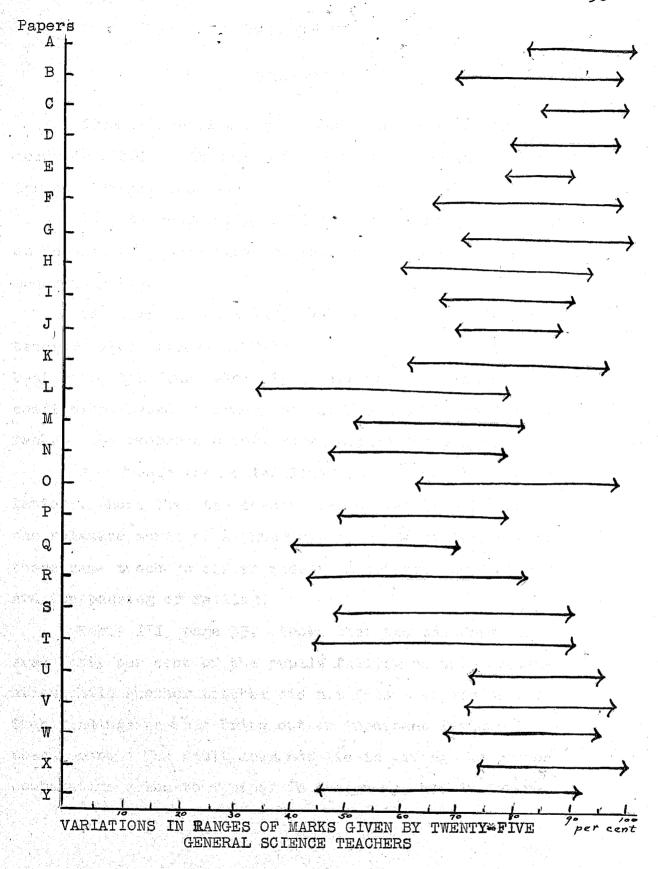| Teacher	Correlation with set of Average Grades	Probable Error	Lowest Passing Grade Indicated	Number of Pupils Failed	Per Cent of Pupils Failed	Number of Pupils Passed	Per Cent of Pupils Passed	Correlation with set of Average Ranks
1	.944	.015	64	7	28%	18	72%	.964
2	.698	.069	70	4	16%	21	84%	.742
3	.955	.012	70	10	40%	15	60%	.889
4	.936	.016	67	7	28%	18	72%	.879
5	.931	.018	69	4	16%	21	84%	.948
6	.916	.022	55	0	0%	25	100%	.891
7	.876	.031	60	4	16%	21	84%	.903
8	.775	.054	65	4	16%	21	84%	.876
9	.947	.014	56	3	12%	22	88%	.898
10	.958	.011	60	3	12%	22	88%	.908
11	.774	.054	69	2	8%	23	92%	.839
12	.844	.038	65	3	12%	22	88%	.865
13	.955	.012	68	10	40%	15	60%	.958
14	.842	.039	62	2	8%	23	92%	.939
15	.618	.083	60	8	32%	17	68%	.848
16	.428	.110	75	3	12%	22	88%	.800
17	.871	.032	70	2	8%	23	92%	.780
18	.904	.024	65	5	20%	20	80%	.854
19	.863	.034	68	0	0%	25	100%	.885
20	.851	.037	65	5	20%	20	80%	.790
21	.937	.016	70	6	24%	19	76%	.925
22	.851	.037	61	5	20%	20	80%	.846
23	.915	.022	70	8	32%	17	68%	.946
24	.937	.016	70	4	16%	21	84%	.872
25	.738	.061	75	3	12%	22	88%	.735

each teacher; the fifth column gives the number of pupils
failed by each teacher, and the sixth column gives the
corresponding per cent of pupils failed; the seventh
column gives the number of pupils passed by each teacher,
and the eighth column gives the corresponding per cent of
pupils passed; the ninth column gives the coefficient
of correlation between the teacher's set of ranks and the
mean or average ranks.

Table IV, page 37, gives the variations in the ranges
of the marks given to the papers by the twenty-five
teachers. The first column lists the papers; the second
column gives the highest grade given to each paper; the
third column gives the lowest grade given to each paper;
and the fourth column gives the range or difference between
the highest and the lowest grades given each paper.

A graph based on the highest and lowest grade given
each paper as found in Table IV, page 37, shows the range
or variation of marks given to the papers by the twenty-
five teachers. In Graph II, page 38, there are twenty-
five lines, one for each paper. The length of each line
is determined by the range or difference in the marks
given that paper by the twenty-five teachers; the shorter
the line the more uniform the teachers are in their grading;
the longer the line the more they vary in their grading.

TABLE IV

SUMMARY OF THE VARIATIONS IN RANGES OF MARKS GIVEN
BY TEACHERS CLASSIFIED ACCORDING TO PAPERS

Paper	Highest Mark Given	Lowest Mark Given	Range or Difference
A	100	82	18
B	98	70	28
C	99	83	16
D	98	81	17
E	90	78	12
F	98	66	32
G	100	70	30
H	93.5	60	33.5
I	90	66	24
J	87.5	69.5	18
K	96	61	35
L	79	33	46
M	81	52.5	28.5
N	78	46.5	31.5
O	98	63	35
P	78	48	30
Q	70	39.5	30.5
R	82	44	38
S	90	47	43
T	90	44	46
U	95	71	24
V	98	71	27
W	95	69	26
X	100	74	26
Y	91	45	46
Correlation with Average Set	.899	.913	

GRAPH II   38

Papers

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y

10   20   30   40   50   60   70   80   90   100
per cent

VARIATIONS IN RANGES OF MARKS GIVEN BY TWENTY-FIVE
GENERAL SCIENCE TEACHERS

CHAPTER VI

CONCLUSIONS

From a careful study of the data presented in the foregoing tables and graphs the writer infers the following outstanding facts:

1. The reliability of the several teachers' marks on the set of general science examination papers is only moderately high.

As shown in Table III, page 35, there are fourteen teachers whose grades correlated less than .900 with the best set, with four below .750. The median correlation coefficient based on grades was .876 and .879 based on ranks. The probable errors were comparatively low.

2. Though the reliability, as revealed by the correlations, shows that the teachers agree fairly well as to the relative merit of a paper within a set of twenty-five, these same teachers differ widely in determining the standard for passing or failing.

Table III, page 35, shows that two teachers considered forty per cent of the pupils failing on this examination while another teacher did not fail a single pupil. Such findings as this bring out an important feature in measurement. The fault does not lie in giving the proper comparative value to a paper in the group, but in deter-

mining the suitable arbitrary mark that decides the success or failure of a student. At this point, the teacher's mark can have a large effect in determining the pupil's confidence in himself and preventing the psychological effects produced by failure. It should be noted here, however, that the difficulty in determining this arbitrary standard is just as difficult in objective tests as in the so-called subjective or essay-type. This is a point too often overlooked by the advocates of the new-type objective test.

The traditional arbitrary 70 or 75 as the passing grade is not so apparent in this study. The passing grades as shown in Table III, page 35, range from 75 to 55 with the median being 67. Because of this, the writer used 65 as the limit of the last group in determining the distribution of grades as shown in Graph I, page 33.

A careful study of Table III, page 35, shows that the per cent of pupils failed or passed does not depend alone on the arbitrary passing grade indicated by the teacher, as one teacher with a passing grade of 70 failed forty per cent of the group of twenty-five while another teacher with the same passing grade, 70, failed only eight per cent of the group of twenty-five. Much depends upon whether a teacher is what is known as a high or a low grader. It is here that the subjective element enters to determine the individual teacher's standard of a "good" paper.

3. It is recognized that an examination should make
a distribution of marks somewhat near the normal curve.
The data in support of this thesis as shown in Graph I,
page 33, reveal a great variation among the twenty-five
teachers in their distribution of marks. The graph of the
average grade bears out its use as the best grade or most
perfect set of marks in that it follows very closely the
normal distribution. A careful study of Graph I, page 33,
shows the wide variation in the judgement of the twenty-
five teachers as to the rating of the twenty-five papers
into the five groups of excellent, good, fair, poor, and
failure.

4. While the correlation of ranks and grades found
in this study was moderately high, attention to Table IV,
page 37, shows that there was a wide range of marks assigned
to the papers. In the case of paper L and paper T the
range was as great as 46 per cent. This difference was
equal to any found by Starch and Elliott in their studies.
However, either the highest set of marks given or the low-
est set would serve as a fairly reliable set of grades for
the twenty-five papers because their correlations with the
average set of grades was .899 and .913 respectively.

It will be noted from Graph II, page 38, that
there was fairly close agreement on the grades assigned to
the better papers. The graph shows by the greater variations

that it was more difficult to measure or determine the value
of the work done by the poor students.

5. Since this study is a comparative study of the sub-
jective grading of general science teachers, some compari-
sons should be made with similar studies previously done.
One might assume that the "scientific method" of approach
to a task might tend to alleviate somewhat the subjective
element in grading essay-type examinations, and that the
correlations would be higher with science teachers. The
median of the correlations found in this study compared with
those found in the Shriner study and the Overholser study
referred to in Chapter II are respectively as follows:

```
Science teachers - .876
Algebra teachers - .946
English teachers - .917
Geometry teachers- .917
```

General Conclusions

To appreciate the value of the results of the study
in the previous section it is well to know some of the
recent thoughts on the subject of measurement. Although
the essay examination has been in existence for hundreds
of years, the amount of research devoted to it is far less
than that on the newer-type, objective tests. What research
there was in the past has been of the negative kind to show
how poor it was in contrast to the new-type test rather
than to show improvement.

In a discussion of the essay tests versus the short-answer tests a comparison of the amount of teacher effort required for each type might be considered. Both tests require eleven hours of expert labor. The essay needing one hour to construct and ten hours to score while the new-type will require ten hours to construct and one hour to score. Five or more studies have been made on the attitudes of the pupils toward each type with the students about evenly divided in their responses. A study made with college students showed that not only was the achievement more permanent, as shown by retests, with the use of the essay examination, but the methods of study were superior. The students' "examination set" was preferable on the essay type.

At its best the essay question calls forth mental processes which are difficult to approach by other means and are extremely desirable objectives in many educational situations. But this form of examination should be restricted to the type of achievement for which it is best suited. Great caution is now used by experts in constructing the essay question so that it calls forth the desired mental processes. The questions are balanced in difficulty and arranged in order starting with the easiest. No choice is permitted among the questions so that student achievement can be comparable.

Because most early studies have proved the unreliabil-

ity of the scoring of the essay type examination, special
emphasis has been given to the various methods of grading.,
Four methods are used:

1. Percentage-passing which gives a false notion of
   fineness of discrimination.

2. Sorting or rating which uses letter grades.

3. A quality scale which does not increase the re-
   liability.

4. Check-list-point score in which each question is
   graded on a series of specifically defined points.
   This is considered the best method and results in
   high reliability.

On the College Entrance Examinations it has been pos-
sible to secure as high a correlation as .990. The scores
are listed for each part of the test by the reader and these
are then totaled and converted to a scale by the technician.
With a close cooperation between reader and technician a
high reliability can be achieved.

Carl Stevason of Notre Dame has devised a "marking
scale" to be used in the conversion of raw scores into
school marks. By the use of this a teacher can use the
check-list-point score method of grading which would produce
higher reliability and then convert the scores to a mark
based on measures of central tendency. To increase the size
of the groups two or more classes of the same teacher or
cooperating teachers may be pooled. Records of grades from
previous semesters may be used in grouping.

Surely by applying the newer procedures in construct-

ing and grading subjective tests and in making a well-
balanced use of both the essay and new-type, objective tests,
a program of measurement and evaluation can be carried on
that will furnish the required data for pupil guidance.

BIBLIOGRAPHY

BIBLIOGRAPHY

Ashburn, Robert, "An Experiment in Essay-Type Questions",
  Journal of Experimental Education, 7: 1-3, September 1938

Bolmeier, E.C., "What's in a Mark?", *School Executive*, 62:25,
  May 1943

Courtis, S. A., "Current Criticisms of Educational Measure-
  ment", *Review of Educational Research*, 8: 545-50, Decem-
  ber 1938

Dunkel, Harold B., "Common Misconception about Testing and
  Evaluation", *School and Society*, May 29, 1943

Greene, Harry A., *Measurement and Evaluation in Secondary
  School*, (New York:Longmans, Green and Company,1942) pp.625

Lawson, Douglas E., "Teachers' Marks--Tragic and Absurd",
  The Educational Forum, 4: 175-179, January 1940

Monroe, Walter S., "Educational Measurement in 1920 and
  1945", *Journal of Educational Research*, January 1945

Oversolser, C.Daniel, "The Significance of the Comparison
  Factor in Measurement", Thesis, Indiana State Teachers
  College, 1940

Remmers, H. H. and Gage, N. L., *Educational Measurement and
  Evaluation*, (New York: Harpers, 1943) pp. 560

Ross, C. C., *Measurement in Todays' Schools*, (New York:
  Prentice-Hall, Inc., 1941) pp.578

Shriner, Walter O., "The Comparison Factor in the Evaluation
  of Examination Papers", *The Teachers College Journal*,
  1: 65-74, January 1930

Shriner, Walter O., "The Reliability of Teachers' Marks",
  The Mathematics Teacher, 17: 426-443, November 1924

Sims, Verner M., "Educational Measurement and Evaluation",
  Journal of Educational Research, September 1944

Sorenson, Herbert, *Psychology in Education*, (New York:
  Mc Graw-Hill Company, 1940) pp. 478

Stalnaker, John M., "Essay Examinations Reliably Read",
  School and Society, 46: 671-672, November 1937

Starch, Daniel and Elliott, Edward C., "Reliability of
    Grading Work in Mathematics", <u>School Review</u>, 21:254-
    259, April 1913

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

Stevason, Carl C., "Simplifying the School Marking Program",
    <u>Journal of Educational Research</u>, April 1945

APPENDIX A

First Semester Test General Science

1. Explain how a plant leaf operates as a food factory.

2. Give several ways in which trees and forests are useful to man.

3. What effect does heat and cold have on matter? Give an example of this scientific principle from everyday experience.

4. Name a chemical element and give its symbol. Give a chemical compound with its formula. How would you classsify the atmosphere? Name a chemical change.

5. Explain how distance to the stars is expressed. Give an example.

6. Choose a heavenly body and discuss it fully.

7. List 5 things in your everyday experience that work by means of air pressure. Which use compressed air?

8. Choose the word from the list that fits the phrase.

   a. the green coloring matter in leaves
   b. the gas that is essential to life and combustion
   c. a group of stars having a special name
   d. an instrument for measuring the pressure of the atmosphere
   e. the process by which water passes into the roots of plants and through the stem
   f. the path followed by a heavenly body
   g. the smallest particle into which all matter can be divided
   h. the sun with the group of planets which revolve around it

osmosis	chlorophyll	constellation
orbit	photosynthesis	solar system
barometer	thermometer	carbon dioxide
compound	meteors	atom
	oxygen	

APPENDIX B

Peru, Indiana
March 15, 1943

My dear fellow science teacher:

I am working on a thesis, "A Compararive Study of Sub-
jective Measurement in General Science", as a part of my
Master's Degree at Indiana State Teachers College. As a
part of this experiment I need to have a set of General
Science examination papers graded by a number of science
teachers that I might make a study of the comparison factor
in educational measurement.

Starch and Elliott made a well-known study on "The
Reliability of the Grading of High School Work in English
and Mathematics". Other studies of teachers' marks have been
made by noted educators. In the Starch and Elliott Study,
one paper was graded by a large number of teachers. The ele-
ment of comparative rank among students could not enter into
such a study.

In this experiment of mine there are a group of twenty-
five students whose papers are to be graded, thus making a
comparison possible. Studies similar to mine have been
made in English and Mathematics by Dr. W.O. Shriner and Mr.
C. Daniel Overholser from Indiana State Teachers College.

I have chosen you as one of the Science teachers of the
State whom I thought would be interested in participating
in this experiment. Your part would be to read the set of
twenty-five papers and give a percentage grade and rank to
each paper. I know these are busy times but it should not
take too long to look over this examination.

Will you please use the enclosed card to let me know,
yes or no, as to your willingness to be a part of this exper-
iment.

I shall be counting on you.

Sincerely yours,

Elizabeth A. Ridenour
R.R. 1   Peru, Ind.

APPENDIX C

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	2	94 %
B	5	89 %
C	1	95 %
D	4	90.5%
E	6	87 %
F	7	83 %
G	3	93.5%
H	11	74 %
I	13	68 %
J	8	78.5%
K	9	78 %
L	18	51.5%
M	16	52.5%
N	16	52.5%
O	5	89 %
P	19	51 %
Q	20	48 %
R	17	56.5%
S	21	47.5%
T	15	64 %
U	12	72.5%
V	10	75.5%
W	12	72.5%
X	6	87 %
Y	14	65.5%
Z		%

INDICATE THE MARK you consider as the lowest passing mark.

64 %

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the study. Such name and address is merely to facilitate in sending the results of the experiment when compiled to the teachers free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	1	98 %
B	3	95 %
C	2	98 %
D	4	94 %
E	7	90 %
F	5	92 %
G	8	89 %
H	9	88 %
I	14	80 %
J	18	85 %
K	11	84 %
L	17	79 %
M	22	65 %
N	21	72 %
O	16	76 %
P	23	62 %
Q	24	60 %
R	12	82 %
S	18	75 %
T	13	81 %
U	6	91 %
V	19	74 %
W	15	79 %
X	20	74 %
Y	25	55 %
Z		%

INDICATE THE MARK you consider as the lowest passing mark. 70 %

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the study. Such name and address is merely to facilitate in sending the results of the experiment when compiled to the teachers free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	3	91 %
B	5	87 %
C	1	94 %
D	2	94 %
E	9	79 %
F	10	77 %
G	4	88 %
H	8	81 %
I	11	77 %
J	15	71 %
K	7	81 %
L	25	37 %
M	16	66 %
N	17	66 %
O	18	66 %
P	23	48 %
Q	24	47 %
R	22	51 %
S	20	57 %
T	21	54 %
U	12	74 %
V	14	72 %
W	13	73 %
X	6	82 %
Y	19	64 %
Z		%

INDICATE
THE MARK you
consider
as the
lowest pass-
ing mark.
70 %

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the
study. Such name and address is merely to facilitate in send-
ing the results of the experiment when compiled to the teachers
free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	1	95 %
B	5	87 %
C	3	93 %
D	4	89 %
E	9	81 %
F	14	77 %
G	2	94 %
H	10	80 %
I	13	79 %
J	15	76 %
K	7	83 %
L	22	63 %
M	21	65 %
N	17	69 %
O	12	80 %
P	19	66 %
Q	24	61 %
R	25	60 %
S	20	66 %
T	18	68 %
U	16	75 %
V	11	80 %
W	8	82 %
X	6	87 %
Y	23	63 %
Z		%

INDICATE THE MARK you consider as the lowest passing mark.
67 %

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the
study. Such name and address is merely to facilitate in send-
ing the results of the experiment when compiled to the teachers
free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	1	98 %
B	4	88 %
C	2	97 %
D	3	95 %
E	10	78 %
F	7	83 %
G	1	98 %
H	11	77 %
I	12	76 %
J	13	74 %
K	5	87 %
L	20	44 %
M	15	70 %
N	17	63 %
O	9	79 %
P	14	71 %
Q	19	52 %
R	18	59 %
S	14	71 %
T	16	69 %
U	8	81 %
V	9	79 %
W	6	85 %
X	1	98 %
Y	9	79 %
Z		%

INDICATE THE MARK you consider as the lowest passing mark. 6.9 %

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the study. Such name and address is merely to facilitate in sending the results of the experiment when compiled to the teachers free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	1	94 %
B	4	89 %
C	3	90 %
D	5	88 %
E	9	80 %
F	14½	77 %
G	2	91 %
H	17	75 %
I	14½	77 %
J	12½	79 %
K	8	84 %
L	23½	60 %
M	20¼	62 %
N	11	81 %
O	10	83 %
P	20¼	62 %
Q	25	55 %
R	20¼	62 %
S	23½	60 %
T	20¼	62 %
U	6	86 %
V	18	72 %
W	12½	79 %
X	7	85 %
Y	16	76 %
Z		%

INDICATE
THE MARK you
consider
as the
lowest pass-
ing mark.
55 %

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the
      study. Such name and address is merely to facilitate in send-
      ing the results of the experiment when compiled to the teachers
      free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	7	85 %
B	12	70 %
C	3	91 %
D	10	81 %
E	9	82 %
F	2	94 %
G	1	95 %
H	6	86 %
I	13	69 %
J	7	85 %
K	9	82 %
L	21	47 %
M	14	67 %
N	18	56 %
O	8	83 %
P	17	60 %
Q	20	50 %
R	17	60 %
S	16	61 %
T	19	51 %
U	10	81 %
V	11	71 %
W	4	89 %
X	5	87 %
Y	15	65 %
Z		%

INDICATE
THE MARK you
consider
as the
lowest pass-
ing mark.
60 %

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the
        study. Such name and address is merely to facilitate in send-
        ing the results of the experiment when compiled to the teachers
        free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	2.5	96 %
B	1	98 %
C	4	94 %
D	2.5	96 %
E	7	85 %
F	5.5	87 %
G	5.5	87 %
H	12.5	81 %
I	12.5	81 %
J	7	74 %
K	12.5	81 %
L	16	75 %
M	18.5	71 %
N	23	58 %
O	15	76 %
P	21	65 %
Q	25	52 %
R	22	61 %
S	20	69 %
T	24	55 %
U	8.5	84 %
V	8.5	84 %
W	13.5	81 %
X	10	83 %
Y	18.5	71 %
Z		%

INDICATE THE MARK you consider as the lowest passing mark.
65 %

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the
study. Such name and address is merely to facilitate in send-
ing the results of the experiment when compiled to the teachers
free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	1	98.5 %
B	5.5	88.5 %
C	4	89 %
D	3	95 %
E	5.5	88.5 %
F	9	81 %
G	2	95.5 %
H	11	77.5 %
I	12	75 %
J	16.5	69.5 %
K	15	72.5 %
L	23	55 %
M	16.5	69.5 %
N	20	63 %
O	8	85 %
P	22	58 %
Q	24	53 %
R	21	61 %
S	18	69 %
T	25	51 %
U	13.5	74 %
V	13.5	74 %
W	10	79.5 %
X	7	85.5 %
Y	19	68 %
Z		%

INDICATE THE MARK you consider as the lowest passing mark.
56 %

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the
study. Such name and address is merely to facilitate in send-
ing the results of the experiment when compiled to the teachers
free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER	
A	3	92 %	
B	3	92 %	
C	3	92 %	
D	5.5	91 %	
E	8	86 %	
F	10	82 %	
G	1	97 %	
H	14.5	77 %	
I	17	66 %	
J	14.5	77 %	
K	9	84 %	
L	25	44 %	
M	16	75 %	
N	18	65 %	
O	11	79 %	
P	23	53 %	
Q	24	47 %	
R	19	64 %	INDICATE
S	20.5	63 %	THE MARKyou
T	22	60 %	consider
U	12.5	78 %	as the
V	12.5	78 %	lowest pass-
W	7	90 %	ing mark.
X	5.5	91 %	60 %
Y	20.5	63 %	
Z		%	

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the
       study. Such name and address is merely to facilitate in send-
       ing the results of the experiment when compiled to the teachers
       free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	3	92 %
B	4	91 %
C	2	87 %
D	3	92 %
E	5	89 %
F	6	88 %
G	2	95 %
H	11	82 %
I	13	75 %
J	9	84 %
K	8	86 %
L	16	65 %
M	14	74 %
N	15	72 %
O	3	92 %
P	17	60 %
Q	12	80 %
R	11	82 %
S	10	83 %
T	10	83 %
U	1	97 %
V	6	88 %
W	3	92 %
X	4	91 %
Y	10	83 %
Z		%

INDICATE
THE MARK you
consider
as the
lowest pass-
ing mark.
    69 %

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the
        study. Such name and address is merely to facilitate in send-
        ing the results of the experiment when compiled to the teachers
        free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	2	98 %
B	9	86 %
C	3	95 %
D	8	$87\frac{1}{2}$ %
E	8	$87\frac{1}{2}$ %
F	7	88 %
G	1	100 %
H	5	93 %
I	13	76 %
J	8	$87\frac{1}{2}$ %
K	3	96 %
L	15	$65\frac{1}{2}$ %
M	11	80 %
N	12	77 %
O	6	90 %
P	17	63 %
Q	18	$40\frac{1}{2}$ %
R	14	$64\frac{1}{2}$ %
S	16	$68\frac{1}{2}$ %
T	15	$65\frac{1}{2}$ %
U	8	$87\frac{1}{2}$ %
V	10	83 %
W	5	93 %
X	6	92 %
Y	4	94 %
Z		%

INDICATE THE MARK you consider as the lowest passing mark.
65 %

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the
study. Such name and address is merely to facilitate in send-
ing the results of the experiment when compiled to the teachers
free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	2	90 %
B	8	82 %
C	3	$89\frac{1}{2}$ %
D	5	$86\frac{1}{2}$ %
E	7	84 %
F	9	76 %
G	1	100 %
H	10	71 %
I	13	68 %
J	13	68 %
K	14	$66\frac{1}{2}$ %
L	22	37 %
M	18	$51\frac{1}{2}$ %
N	20	$46\frac{1}{2}$ %
O	12	$68\frac{1}{2}$ %
P	19	$49\frac{1}{2}$ %
Q	21	$39\frac{1}{2}$ %
R	17	$55\frac{1}{2}$ %
S	16	59 %
T	16	59 %
U	11	70 %
V	6	$85\frac{1}{2}$ %
W	7	84 %
X	4	$88\frac{1}{2}$ %
Y	15	61 %
Z		%

INDICATE THE MARK you consider as the lowest passing mark.

_68 %_

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the
study.   Such name and address is merely to facilitate in send-
ing the results of the experiment when compiled to the teachers
free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	1	100 %
B	9	88 %
C	2	99 %
D	3	98 %
E	5	92 %
F	8	89 %
G	4	97 %
H	20	48 %
I	11	82 %
J	14	74 %
K	7	90 %
L	18	62 %
M	13	76 %
N	15	69 %
O	9	88 %
P	16	66 %
Q	19	54 %
R	14	74 %
S	15	69 %
T	17	64 %
U	8	89 %
V	12	80 %
W	10	84 %
X	6	91 %
Y	13	76 %
Z		%

INDICATE THE MARK you consider as the lowest passing mark.

62 %

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the
study. Such name and address is merely to facilitate in send-
ing the results of the experiment when compiled to the teachers
free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	1	93 %
B	7	74 %
C	3	86 %
D	2	90 %
E	5	79 %
F	13	66 %
G	10	70 %
H	16	60 %
I	10	70 %
J	5	79 %
K	15	61 %
L	24	33 %
M	16	60 %
N	19	56 %
O	14	63 %
P	23	36 %
Q	25	29 %
R	21	49 %
S	20	54 %
T	18	58 %
U	9	71 %
V	4	81 %
W	12	69 %
X	7	74 %
Y	22	45 %
Z		%

INDICATE THE MARK you consider as the lowest passing mark. 60 %

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the study. Such name and address is merely to facilitate in sending the results of the experiment when compiled to the teachers free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	1	97 %
B	2	96 %
C	13	87 %
D	4	92 %
E	7	90 %
F	3	93 %
G	10	90 %
H	14	85 %
I	8	90 %
J	16	80 %
K	6	91 %
L	22	75 %
M	17	80 %
N	23	73 %
O	19	79 %
P	25	70 %
Q	20	78 %
R	24	73 %
S	15	83 %
T	21	75 %
U	5	92 %
V	12	88 %
W	9	90 %
X	11	89 %
Y	18	80 %
Z		%

INDICATE THE MARK you consider as the lowest passing mark.

75 %

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the
       study. Such name and address is merely to facilitate in send-
       ing the results of the experiment when compiled to the teachers
       free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	3	95 %
B	9	85 %
C	3	95 %
D	6	91 %
E	8	88 %
F	11	81 %
G	2	97 %
H	7	90 %
I	10	83 %
J	9	85 %
K	4	93 %
L	16	53 %
M	11	81 %
N	12	78 %
O	1	98 %
P	14	72 %
Q	15	63 %
R	12	78 %
S	7	90 %
T	13	74 %
U	8	88 %
V	3	95 %
W	3	95 %
X	2	97 %
Y	5	91 %
Z		%

INDICATE THE MARK you consider as the lowest passing mark.

70 %

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the
study. Such name and address is merely to facilitate in send-
ing the results of the experiment when compiled to the teachers
free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	1	94 %
B	4	88 %
C	1	94 %
D	3	91 %
E	4	88 %
F	1	94 %
G	1	94 %
H	6	82 %
I	4	88 %
J	9	79 %
K	7	81 %
L	17	52 %
M	10	78 %
N	12	70 %
O	4	88 %
P	14	64 %
Q	16	55 %
R	13	66 %
S	14	64 %
T	15	59 %
U	8	80 %
V	5	84 %
W	11	74 %
X	2	92 %
Y	10	78 %
Z		%

INDICATE
THE MARK you
consider
as the
lowest pass-
ing mark.
65 %

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the
study. Such name and address is merely to facilitate in send-
ing the results of the experiment when compiled to the teachers
free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	1	98 %
B	4	92 %
C	2	94 %
D	4	92 %
E	8	85 %
F	11	81 %
G	3	93 %
H	3	93 %
I	6	90 %
J	10	82 %
K	12	79 %
L	16	74 %
M	15	75 %
N	13	78 %
O	4	92 %
P	18	68 %
Q	17	70 %
R	14	76 %
S	17	70 %
T	16	74 %
U	2	94 %
V	5	91 %
W	7	88 %
X	7	88 %
Y	9	83 %
Z		%

INDICATE THE MARK you consider as the lowest passing mark.

68 %

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the study. Such name and address is merely to facilitate in sending the results of the experiment when compiled to the teachers free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	3	93 %
B	5	89 %
C	6	83 %
D	1	97 %
E	2	94 %
F	4	91 %
G	4	91 %
H	6	83 %
I	7	78 %
J	8	76 %
K	7	80 %
L	15	50 %
M	8	76 %
N	13	55 %
O	5	89 %
P	10	65 %
Q	14	52 %
R	10	65 %
S	11	62 %
T	12	60 %
U	8	76 %
V	7	80 %
W	9	73 %
X	4	91 %
Y	5	89 %
Z		%

INDICATE THE MARK you consider as the lowest passing mark.
65 %

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the study. Such name and address is merely to facilitate in sending the results of the experiment when compiled to the teachers free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	3	90 %
B	1	94 %
C	3	90 %
D	2	92 %
E	4	86 %
F	9	75 %
G	3	90 %
H	9	75 %
I	11	73 %
J	10	74 %
K	6	83 %
L	18	49 %
M	12	72 %
N	13	71 %
O	8	76 %
P	16	54 %
Q	17	50 %
R	17	50 %
S	14	67 %
T	15	64 %
U	7	80 %
V	4	86 %
W	8	76 %
X	5	85 %
Y	7	80 %
Z		%

INDICATE THE MARKyou consider as the lowest passing mark.
70 %

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the study. Such name and address is merely to facilitate in sending the results of the experiment when compiled to the teachers free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	6	82 %
B	1	95 %
C	5	85 %
D	2	93 %
E	3	91 %
F	11	70 %
G	4	90 %
H	12	65 %
I	5	85 %
J	6	82 %
K	4	90 %
L	14	49 %
M	13	61 %
N	9	76 %
O	10	73 %
P	15	48 %
Q	16	44 %
R	16	44 %
S	7	79 %
T	16	44 %
U	5	85 %
V	8	78 %
W	8	78 %
X	4	90 %
Y	9	76 %
Z		%

INDICATE THE MARK you consider as the lowest passing mark.
61 %

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the study. Such name and address is merely to facilitate in sending the results of the experiment when compiled to the teachers free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	1	100 %
B	4	94 %
C	3	96 %
D	5	90 %
E	9	84 %
F	9	84 %
G	2	98 %
H	10	81 %
I	11	79 %
J	8	85 %
K	6	89 %
L	21	54 %
M	15	69 %
N	13	75 %
O	7	88 %
P	18	60 %
Q	20	55 %
R	17	61 %
S	19	58 %
T	16	67 %
U	12	77 %
V	13	75 %
W	14	73 %
X	6	89 %
Y	20	55 %
Z		%

INDICATE THE MARK you consider as the lowest passing mark.

70 %

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the study. Such name and address is merely to facilitate in sending the results of the experiment when compiled to the teachers free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	1.5	100 %
B	4	96 %
C	4	96 %
D	6.5	95 %
E	10	91 %
F	6.5	95 %
G	4	96 %
H	11.5	88 %
I	15.5	81 %
J	17	80 %
K	8	93 %
L	25	62 %
M	19	77 %
N	18	78 %
O	9	92 %
P	22	67 %
Q	23	65 %
R	20	73 %
S	21	70 %
T	24	64 %
U	13	86 %
V	14	82 %
W	15.5	81 %
X	1.5	100 %
Y	11.5	88 %
Z		%

INDICATE THE MARK you consider as the lowest passing mark. _70 %_

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the study. Such name and address is merely to facilitate in sending the results of the experiment when compiled to the teachers free of charge.

RECORD OF GRADES

STUDENT	RANK OF PAPER	% GRADE GIVEN PAPER
A	6	95 %
B	9.5	90 %
C	6	95 %
D	3	98 %
E	9.5	90 %
F	3	98 %
G	15	85 %
H	17	82 %
I	9.5	90 %
J	15	85 %
K	13	87 %
L	21.5	75 %
M	21.5	75 %
N	23	72 %
O	18	80 %
P	19.5	78 %
Q	24.5	70 %
R	19.5	78 %
S	15	85 %
T	9.5	90 %
U	6	95 %
V	3	98 %
W	12	88 %
X	1	100 %
Y	24.5	70 %
Z		%

INDICATE
THE MARK you
consider
as the
lowest pass-
ing mark.

_75 %_

Graded by:

Name _____

Address _____

Note: The name of teacher grading the paper will not be used in the
study. Such name and address is merely to facilitate in send-
ing the results of the experiment when compiled to the teachers
free of charge.