

RELATION OF ENGLISH COMPOSITIONS WRITTEN DURING SPRING SEMESTER
OF 1951 BY SOPHOMORES ENROLLED AT SOUTHWEST TEXAS STATE
TEACHERS COLLEGE TO COOPERATIVE ENGLISH TEST TAKEN
DURING SPRING SEMESTER OF 1951 BY THE SAME GROUP

Approved:

D. F. Motlow

J. Lloyd Rogers

L. N. Wright

Approved:

Claude Ellorff

Dean, Graduate Studies

RELATION OF ENGLISH COMPOSITIONS WRITTEN DURING SPRING SEMESTER
OF 1951 BY SOPHOMORES ENROLLED AT SOUTHWEST TEXAS STATE
TEACHERS COLLEGE TO COOPERATIVE ENGLISH TEST TAKEN
DURING SPRING SEMESTER OF 1951 BY THE SAME GROUP

THESIS

Presented to the Graduate Council of
Southwest Texas State Teachers College
in Partial Fulfillment of
the Requirements

For the Degree of

MASTER OF ARTS

By

Paula Kae LaForge, B. S.
(Marble Falls, Texas)

San Marcos, Texas

August, 1952

ACKNOWLEDGMENTS

The writer wishes to express her sincere gratitude to Dr. David F. Votaw, chairman of the committee, for his advice and guidance during the preparation of this study.

Special acknowledgment is also made of the assistance of Dr. L. N. Wright and of Dr. J. Llc , who served as members of the committee.

Paula Kae LaForge

San Marcos, Texas

June, 1952

TABLE OF CONTENTS

Chapter	Page
I. THE PROBLEM DEFINED	1
Statement of the Problem	1
Importance of the Study	3
Limitations of the Study	3
Definitions of Terms	3
Adequacy of Evaluation	5
Standardized English test	5
Written Compositions	5
Aids for Achieving Accurate and Effective Evaluation	13
Evaluation of Southwest Texas State Teachers College Sophomore Written Compositions	15
The General Plan of Evaluation	16
Adequacy of Evaluation of Southwest Texas State Teachers College Papers	19
II. TREATMENT OF THE DATA	25
Description of the Data	25
Methods Used in Comparing the Data	26
Results of Computations	26
Composition Ratings as Criteria	26
Relationships between the Two Sets of Data	28
Investigation of Causes	37
The Cutoff Point	40
III. SUMMARY AND CONCLUSIONS	43
APPENDIX	44
BIBLIOGRAPHY	51

LIST OF TABLES

Table	Page
I. Composition-Rating Chart	17
II. Interpretation of Composition Ratings	18
III. Correlation Coefficients for Written Compositions and Cooperative English Tests	29
IV. Effectiveness of Expression (Actual Numbers)	30
V. Effectiveness of Expression (Percentages)	31
VI. Reading Comprehension (Actual Numbers)	32
VII. Reading Comprehension (Percentages)	33
VIII. Total English (Actual Numbers)	34
IX. Total English (Percentages)	35
X. Correlation of Composition Ratings and Effectiveness of Expression	36

CHAPTER I

THE PROBLEM DEFINED

Written composition has been used for many years as a basic means for determination of a student's ability to use the English language effectively. The standardized test, although of a more recent origin, is also widely used for the same purpose. The comparisons made between these two bases in this study are exploratory and very limited.

Statement of the Problem

In any college or university, there are those students who have not acquired an adequate working ability to do written composition by the time of completion of the first English composition courses which are usually required during the first year of college. Hence, the need arises for remedial work. The next step is to discover which students should be required to do additional work in English composition. As a part of the solution to this problem, the English department of Southwest Texas State Teachers College set up a tentative plan in the form of a composition-evaluating program for evaluating its sophomores' ability to do written composition.

Because of the great amount of time and effort required to carry out an annual program of composition evaluation for

sophomores, the question arose as to the possibility of there being some substitute procedure which would do the same job effectively, but which would require a lesser amount of time and effort.

Since an objective testing program which includes all Southwest Texas State Teachers College sophomores is also administered annually during the spring semester, the question was raised as to the use of the English tests included in this program as a means of evaluating sophomore writing ability.

It was the purpose of this study to investigate this possibility (1) by comparing the level of performance in written composition of each sophomore with his score made on a standardized test in English, and (2) by using the written compositions as a criterion for determining adequate ability in order to investigate further by establishing a cutoff point on the standardized test score-scale to mark the minimum adequacy of ability to do written composition.

The comparisons were made primarily by correlating the composition ratings with test scores made on the whole, the parts, and groups of the parts of the standardized English test.

Importance of the Study

This study is the first to be made of the present sophomore English evaluating program in this college. The spring semester of 1951 was the first semester during which sophomore students were required to write compositions within an organized composition-evaluating program as a means of evaluation of ability; that is, it was the first time they were required to write using criteria set up by the English department as a whole as a means of evaluation. However, the standardized English test has been given to several previous classes of sophomore students. This study is an exploratory one which serves possibly as a basis for further investigation.

Limitations of the Study

Since the evaluation program is a relatively new one and is experimental in nature, this study is limited to information which could be taken from work done in the spring semester of 1951.

Definitions of Terms

Composition.--The two primary methods of communication, speaking and writing, can be placed in one category: composition. Composition implies composing, that is, putting

thoughts together and communicating them by speaking or writing so as to be understood by others.¹ Written composition is the only phase which will be dealt with in this study. Harry Shaw says, "The whole process of written composition consists of two steps: . . . the first step is thinking; the second is writing. These two are inseparably linked."²

Evaluation.--When the value or amount of thinking and writing which have been produced by an individual in writing a composition has been ascertained, the composition is said to have been evaluated.

Adequacy of evaluation.--If the evaluation of a paper or group of papers is sufficient for the fulfillment of its intended purpose, then the evaluation is said to be adequate.

T-score.--Any distribution of measures whose mean is fifty and whose standard deviation is ten is said to be in the form of T-scores.

¹Henry Seidel Canby and John Baker Opdycke, Elements of Composition for Secondary Schools, p. 1.

²Harry Shaw, Complete Course in Freshman English, p. 4.

Adequacy of Evaluation

In this section of the study, the attempt is made to determine whether the evaluation of the standardized English tests and the written compositions was adequate.

Standardized English Test

The standardized English test used in this case was entirely objective, and the answer sheets were scored on a rights-minus-wrongs basis. Sufficient tests have been made by the authors of the test to prove its appropriate validity and reliability. Therefore, it is safe to assume that the evaluation of the standardized English test was adequate.

Written Compositions

The effectiveness of the evaluation of the compositions is not nearly so objective as that of the standardized test and must be considered from several viewpoints other than that of merely scoring answer sheets right or wrong as is done in the case of the standardized tests. The impossibility of judging composition accurately as right or wrong is a well-known fact; hence, the following paragraphs will be devoted to an investigation of the evaluation of written English composition and then, more specifically, to an investigation of the evaluation of the compositions used in this study.

The writer's purposes in making this investigation of the evaluation of written English composition may be defined by the following statements:

1. To discover what characteristics of composition are evaluated
2. To determine some characteristics of an effective evaluator
3. To review some prevalent methods of composition evaluation
4. To compile some aids for achieving accurate and effective evaluation of written composition

Characteristics of composition which should be evaluated.

The characteristics of a paper which might be considered in evaluating its composition are innumerable. Probably three which are best known are the old favorites, unity, coherence, and emphasis. There are many others, however, which may or may not be included when one is speaking of the aforementioned characteristics.

Joseph M. Bachelor and Ralph L. Henry list in their chart for checking a composition such things as (1) choice of subject, (2) unity, (3) coherence, (4) emphasis, (5) words, (6) punctuation, and (7) mechanical details.³

³Joseph M. Bachelor and Ralph L. Henry, American Thinking and Writing, p. 344.

Ida Jewett says, "Content, organization, form--but the greatest of these is content," and Charles S. Pendleton states, "An instructor does well who reads a theme chiefly to get the thing which the pupil chiefly intended to put into it."⁴ However, many times the student does not know himself what he intended to put into the paper.

One popular book on composition stresses the appearance of a paper.⁵ It is that author's opinion that a student may not be capable of writing beautifully, but certainly he can write legibly. Almost any English teacher will agree that appearance is important, especially if he recalls innumerable times when he has sat with a paper before him trying desperately to distinguish between expressions such as "good eggs" and "gold dogs."

George Henry feels that the primary attribute of composition is sincerity on the part of the writer.⁶ Encouragement of sincerity helps the student to know himself better and, therefore, to develop his potentialities as an individual.

⁴Lucile Turner, "Theories and Practices in Evaluation and Correction of Oral and Written Composition," Peabody Journal of Education, XIX (March, 1942), 267.

⁵Thomas E. Rankin, Clarence D. Thorpe, and Melvin T. Solve, College Composition, pp. 755-760.

⁶George Henry, "How to Get Interesting Themes," The English Journal, XXIII (September, 1944), 348.

Canby says that good writing presents three specific problems: (1) straight thinking, (2) adequate expressions, and (3) good form.⁷ When these three characteristics of acceptable composition have been mastered to a reasonable degree, then Canby thinks that the student may begin to think about developing style.

Still another criterion for evaluation is used by Lucia B. Mirrielees. She divides the marking into two parts: content and form. By content is meant the way that the subject matter has been selected, realized, organized, and presented. The technical errors which are made constitute errors in form.⁸

George S. Wykoff says that the purpose of composition is the achievement of correct, clear, and effective writing.⁹ A paper should be marked according to its purposes; hence, correctness, clearness, and general effect of the manuscript are characteristics to be evaluated. Wykoff continues by saying that the teacher should partly judge the composition as to

⁷Henry Seidel Canby, English Composition in Theory and in Practice, pp. xiii-xiv.

⁸Lucia B. Mirrielees, Teaching Composition and Literature, pp. 35-61.

⁹George S. Wykoff, "Suggestions for the Reading of Themes," College English, XXI (January, 1950), 210.

its adaptation to the reader and that it is very important that the teacher make use of a set of grading standards for technical errors when marking a composition.

The preceding paragraphs mention only a few characteristics which may or may not, depending upon the marker, be taken into consideration when reading a composition for evaluation purposes. Innumerable charts for marking mechanical details, in which are listed the many common errors which occur, are readily available. As for the appearance, legibility, sincerity, straight thinking, adequate expression, general effect of the manuscript, and all the many other characteristics which are intangible, the privilege to discriminate is left largely to the discretion of the evaluator. But, in any evaluation, especially when a number of papers are involved, the characteristics to be evaluated should be definitely set down so as to leave as little leeway as possible to the discretion of the marker.

Characteristics of an effective evaluator.--The qualities which a composition may possess, whether the student has displayed a moronic point of view or the insight of a genius, matter little if the evaluator is not competent. Competency may include any or all of a number of things.

Percy Isaac Reed and Elizabeth Frost Reed state that the reader must be academically competent to judge and, above all

else, he must be frank and unprejudiced.¹⁰ Raymond W. Pence says, "After spending his time writing a composition, a student is entitled to the very best in the way of criticism of it that a teacher can give."¹¹ From this point of view, it could be said that the evaluator should possess the willingness to give his best in time, training, and attitude.

Objectivity on the part of the evaluator is of utmost importance. Wykoff states that the teacher must not let his like or dislike of the chosen subject or his agreement or disagreement with ideas presented influence the mark which he gives a paper.¹² Objectivity of attitude toward students is even more important, of course, than toward subject matter.

The foregoing paragraphs may be summarized by the following statement: To evaluate composition adequately, one should be (1) academically competent, (2) objective in attitude, and (3) willing to give sincere attention to the task at hand. These are characteristics which the writer believes to be among the most important requirements for effective evaluation of a composition.

¹⁰Percy Isaac Reed and Elizabeth Frost Reed, Applied Composition, p. 343.

¹¹Raymond W. Pence, College Composition, p. ix.

¹²George S. Wykoff, loc. cit.

Prevalent methods of composition evaluation.--In performing any task, there is usually a choice of methods which may be employed. For evaluation of written English composition, perhaps the one most widely used is that of assigning letter marks, such as A (superior), B (above average), C (average), D (below average), and F (failure). A second method which is widely used is that of ranking papers according to a given set of numbers, such as 1 (superior), 2 (satisfactory), and 3 (unsatisfactory).¹³

In some cases the percentage basis is still used. The writer is inclined to agree with Lucia B. Mirrielees, however, when she says, "You know and I know, of course, the absurdity of marking one bit of human experience 85% and another 87%."¹⁴ She says that written composition is in reality nothing but human experience, whether it be of the physical, mental, or spiritual form. Perhaps, when observed in the light of measuring human experience, the whole system of evaluating composition seems to be of little value, and the percentage basis seems merely to be a bit more inadequate than others because it does not afford as much flexibility as some other methods.

¹³Charles Swain Thomas, The Teaching of English, pp. 51-70.

¹⁴Lucia B. Mirrielees, op. cit., p. 47.

To increase flexibility of marks to some degree, Mirrielees has devised a method of marking composition whereby two marks are given to a composition, that is, $\frac{A}{C}$, $\frac{D}{B}$, $\frac{B}{A}$, $\frac{C}{B}$, in which the content credit is recorded by the top mark and the credit given for form by the bottom mark.¹⁵ The author of this system claims a twofold advantage for it. First, the grader is assisted in obtaining a broader outlook of the paper. If he gives the student a grade on both content and form, he is forced to judge the paper from at least two points of view. The second advantage which might be mentioned is the fact that this method is of value to the student himself because, by receiving his mark in two divisions, he can more readily see his weaknesses and his strong points.

Another method for determining marks is that based on arithmetic computation involving the number of words contained in the composition and the number of errors made.¹⁶ This method was derived by Helen Bosworth, West Springfield, Massachusetts, and it is a variation of Lucia B. Mirrielees' method.

For more objective evaluation, several composition scales have been developed. Mirrielees gives the following discussion of their relative value:

¹⁵Ibid., pp. 35-61.

¹⁶Ibid., p. 495.

. . . the chief difficulty of these scales is that of applying them. One pupil differs so greatly from another in vocabulary, style of writing, personality, that it is difficult to compare a pupil's paper with a printed model and reach a conclusion. Teachers can train themselves to use composition scales, but the average teacher with a full schedule will find that while scales dealing with mechanics are invaluable, scales dealing with composition content are exceedingly difficult and time-taking. Three of the well-known composition scales are these: Huddelson's English Composition Scale (World Book Company, Yonkers, New York); Nassau County Supplement to the Hillegas Scale for Composition (World Book Company, Yonkers, New York); and the letter scales called Scales for Measuring Special Types of Compositions (World Book Company, Yonkers, New York).¹⁷

Aids for Achieving Accurate and Effective Evaluation

The following list of aids for achieving accurate and effective evaluation of written composition was compiled from an article by Paul B. Diederich:

1. Papers must not be marked for content alone
2. The topic must be within the scope of the student but not too easy
3. The form of writing must be one with which the students are familiar
4. The topic should have been selected some time ahead of the writing period
5. The composition must be written in class to prevent outside assistance

¹⁷Ibid., p. 49.

6. At least two essays on different topics are needed for reliable measure
7. All students should write on the same assignment
8. Sufficient time must be allowed
9. The papers must be judged according to formulated criteria made in advance¹⁸

An essential requisite for accurate evaluation is given by Austin J. App, who states, "When I grade the papers of my students my scale is not the norm of perfection of the ages, of Shakespeare and Dante, but of the classroom."¹⁹

There are probably as many different sets of criteria for composition evaluation as there are English instructors. Robinson H. Shipherd says,

The reader must have the most fair and rational understanding of the evaluation criterion as it is reasonably possible to have. It makes astonishingly little difference what those methods are, so they be mainly reasonable and consistent.²⁰

It may be noted that each of the aids listed in the foregoing discussion is basic to composition marking. It is impossible to say, though, that a composition has been

¹⁸Paul B. Diederich, "The Measurement of Skill in Writing," School Review, LIV (December, 1946), 588.

¹⁹Austin J. App, "Old Ben Jonson on Grading Compositions," College English, XIV (February, 1943), 318.

²⁰Robinson H. Shipherd, The Fine Art of Writing for Those Who Teach It, p. iii.

evaluated adequately if the list has been adhered to, or that a composition has not received the necessary attention if these suggestions are not followed. The list is flexible. Any set of criteria may be used, but that set must be followed precisely by the marker once it is set up if the evaluation is to be valid.

Perhaps the evaluation of a written composition, if it is effectively done, is the most difficult and disconcerting task that a teacher must undertake. There are many aspects of the process which are intangible and the marking of which is left to the discretion of the manuscript reader. However, if the evaluations of compositions are to be comparable, there is one precedent which must not be violated. The evaluations must be based on one common set of criteria.

Evaluation of Southwest Texas State Teachers College Sophomore Written Compositions

The purpose of the evaluation of the set of sophomore written compositions has previously been stated in the beginning of the study as follows: to discover which students have not acquired an adequate ability to do English written composition and, therefore, need additional work in English. A complete list of the students' test scores and composition ratings is given in the Appendix of this study.

The General Plan of Evaluation

Each student who was classified as a sophomore and who was currently enrolled in an English class during the spring semester of 1951 was required to write a composition. The work was done in the classroom, and the students were given a two-hour period in which to write their compositions. They were instructed to write on some subject pertaining to their major field of study while in college.

The form presented in Table I was used as a criterion for the evaluation of the compositions. Each student's paper was rated according to the composition-rating chart by a faculty member from the student's major department. Then the paper was rated by a member of the English department. If, after these two ratings, the paper received combined ratings of (1) and (1) or (1) and (2), the paper was not rated again, and it was assumed that the student's ability to do written composition was adequate. If, after two ratings, the paper did not receive either of the above rating combinations of (1) and (1) or (1) and (2), it was sent to a third reader, usually a person from a department other than the student's major department or the English department. In that manner it was given a third rating.

If any paper received as low as even a single total grade rating of (3) from any one of the three readers, that

TABLE I
COMPOSITION-RATING CHART

WRITTEN ENGLISH GRADE SHEET

STUDENT'S NAME _____ DEPARTMENT _____

GRADE

SYMBOLS: 1 equals "clearly satisfactory"

2 indicates some doubt on the part of the grader

3 equals "clearly unsatisfactory"

Grader	Spell.	Leg.	Punct.	Sent. Struct.	Gram.	Organ.	Gen. Effect	Grade
--------	--------	------	--------	------------------	-------	--------	----------------	-------

Spelling, legibility, and punctuation are relatively of minor importance in grading, with grammar, sentence structure, organization, and general effectiveness in developing the topic of major importance.

student's ability to do written composition was said to be inadequate according to the ratings; and he was advised to do some remedial work, perhaps in the form of an advanced composition or grammar course, depending upon his individual needs. Those papers which were rated three times and did not receive any (3)'s were designated as satisfactory work.

For greater ease in comparing the composition ratings with the test scores, all the combinations of composition ratings were ascertained and then rated according to their various numerical averages. Table II explains the composition ratings 0 through 6 which will be used throughout this study instead of the original (1), (2), and (3).

TABLE II
INTERPRETATION OF COMPOSITION RATINGS

Occurring Combinations	Numerical Averages	Composition Ratings
1 and 1	1.00	6
1 and 2	1.50	5
2, 2, and 1	1.67	4
2, 2, and 2	2.00	3
2, 2, and 3	2.33	2
3 and 2	2.50	1
3 and 3	3.00	0

Adequacy of Evaluation of Southwest Texas State
Teachers College Papers

As stated in a previous section, the adequacy of evaluation of a written composition depends upon the fulfillment of the intended purpose of the evaluation, which, in this case, was to discover whether or not the sophomore students could produce acceptable written composition. However, the fulfillment of this purpose depends upon several factors. Four which were considered the most important by the writer have been previously investigated in this study in the section dealing with "Adequacy of Evaluation." They refer to (1) characteristics of composition which are evaluated, (2) characteristics of an effective evaluator, (3) methods of evaluation, and (4) aids to effective evaluation of composition. It would appear, therefore, that if the evaluation of the Southwest Texas State Teachers College papers fulfilled the criteria set up by these four factors, then the evaluation was adequate.

In a previous section, numerous characteristics of compositions which could be evaluated were listed. The conclusion was that it matters little which characteristics are evaluated; the most important point is that the characteristics be set up definitely, so as to leave as little leeway as possible to the discretion of the marker. In the case of the sophomore papers, the characteristics were

listed definitely on the composition-rating chart to be rated (1), (2), or (3). The evaluators knew which ones were to be evaluated; therefore, the evaluation fulfilled the criteria set up for the first factor.

The second factor dealt with the characteristics of an effective evaluator. The criteria previously stated were that the evaluator should be (1) academically competent, (2) objective in attitude, and (3) willing to give sincere attention to the task at hand. The evaluators of the sophomore papers were all members of the college faculty. Therefore, it was assumed that they were academically competent. The second criterion, objectivity in attitude, is more intangible and therefore more difficult to judge whether it was fulfilled sufficiently. Differences in training of faculty members might reasonably be expected to produce variations in their approximation to this criterion.

As was previously stated, the teacher must not let his like or dislike of the chosen subject or his agreement or disagreement with ideas presented influence the mark which he gives a paper. Like or dislike of subject matter of the compositions could probably be discarded as insignificant in the case of the sophomore papers. The first rating was given by evaluators who marked papers dealing with their

respective departments, which subjects they probably liked; the second ratings were given by a member of the English department who, because of his knowledge of the variety and nature of the composition-evaluating program and because of his interest in composition, would not be likely to discriminate because of subject matter. The third rating, for which the paper was sent to an academic department other than that of the student's major or the English department, would probably be the most vulnerable to this discrimination. However, it is possible that this aspect of objectivity is relatively insignificant in the over-all results.

A possible hindrance to objectivity of attitude toward students might have been the fact that the papers were identified by the students' names. This practice is especially dangerous in this case because students' papers were marked by faculty members from the students' major fields of study. The idea of preconceived judgment of students' abilities could have been entirely cast out had the papers been unidentifiable by the evaluators.

From this evidence it was concluded that the possibility exists that the attitude of the evaluators may or may not have been objective; but by the ethics of good teaching, it has been presumed that they were objective to the best of their abilities.

The third criterion for an effective evaluator is that he be willing to give sincere attention to the task at hand. This is also a matter of ethics, and it has been assumed that each evaluator gave each paper sufficient study and consideration.

The third factor for fulfilling the purpose of adequate evaluation was method of evaluation. The method used in the evaluation of the sophomore papers has been discussed previously; it consisted of a composition-rating chart on which the compositions' characteristics were rated (1) clearly satisfactory, (2) doubtfully satisfactory, and (3) clearly unsatisfactory. This was one of the methods discussed in the section on methods of evaluation.

One discrepancy in regard to this method has become prevalent in the case of the evaluation of the sophomore papers. This arose from the fact that the evaluation was carried out by a number of evaluators. This was done to conserve time; however, the fact remains that when a sizable group of papers is marked by one person, it is the tendency of the marker to evaluate the papers along the line of the normal probability distribution. Normally, this method is correct and the only one which should be used. However, when a group of papers constitutes only a small part of a large group, such as was the case in this evaluation, where each

instructor was given only a few papers to mark, this method is to some degree invalid. For example, the evaluator might be given only three papers which might possibly be the three highest ranking papers in the entire group. Yet, knowing of no lower ranking papers, the marker might rate the three highest ones as (1), (2), and (3) when in reality, if rated with the entire group, the papers should have all received ratings of (1). The discrepancy lies in the fact that the evaluator would be deprived of the all-important normative reference by which to determine ratings of (1), (2), or (3) and, therefore, be compelled to resort to so-called standards of his own. Such standards are known to differ widely from marker to marker.

Edmiston and Gingerich made a study similar to this one, however, in which the results of their evaluation were much the same as those in this study; yet their papers were marked by the Huddleson Typical Composition Scale by one evaluator.²¹ Perhaps then the discrepancy involved when several evaluators are used was not as great in this case as one might be inclined to think.

The fourth factor investigated previously dealt with aids in achieving accurate and effective evaluation. The

²¹R. W. Edmiston and C. N. Gingerich, "The Relation of Factors of English Usage to Composition," Journal of Educational Research, XXXVI (December, 1942), 269-271.

instructor was given only a few papers to mark, this method is to some degree invalid. For example, the evaluator might be given only three papers which might possibly be the three highest ranking papers in the entire group. Yet, knowing of no lower ranking papers, the marker might rate the three highest ones as (1), (2), and (3) when in reality, if rated with the entire group, the papers should have all received ratings of (1). The discrepancy lies in the fact that the evaluator would be deprived of the all-important normative reference by which to determine ratings of (1), (2), or (3) and, therefore, be compelled to resort to so-called standards of his own. Such standards are known to differ widely from marker to marker.

Edmiston and Gingerich made a study similar to this one, however, in which the results of their evaluation were much the same as those in this study; yet their papers were marked by the Huddleson Typical Composition Scale by one evaluator.²¹ Perhaps then the discrepancy involved when several evaluators are used was not as great in this case as one might be inclined to think.

The fourth factor investigated previously dealt with aids in achieving accurate and effective evaluation. The

²¹R. W. Edmiston and C. N. Gingerich, "The Relation of Factors of English Usage to Composition," Journal of Educational Research, XXXVI (December, 1942), 269-271.

Southwest Texas State Teachers College plan fulfilled all the criteria listed except the one stating that for the sake of reliability the students should be rated on at least two compositions dealing with different topics. Because of the experimental nature of this program, the students were only required to write one composition instead of two.

Probably the most important of the aids listed is the requirement that all evaluations be based on one set of criteria if they are to be comparable. This criterion, though the simplest minimum, was set up for the sophomore papers in the composition-rating chart.

By this investigation of the factors involved in adequate composition evaluation, it has been shown that the evaluation of the sophomore written compositions at Southwest Texas State Teachers College was not entirely adequate. As long as any evaluation is subjective, there will be discrepancy to some degree.

CHAPTER II

TREATMENT OF THE DATA

Description of the Data

This study deals with the relationships of two sets of data, namely, (1) a group of written compositions and (2) a group of Cooperative English Test C₂, Form Y papers. The Cooperative English Test includes the following parts:

- I. Reading Comprehension
 - A. Vocabulary
 - B. Reading
- II. Mechanics of Expression
 - A. Grammatical Usage
 - B. Punctuation and Capitalization
 - C. Spelling
- III. Effectiveness of Expression
 - A. Sentence Structure and Style
 - B. Active Vocabulary
 - C. Organization

The students were asked to write their compositions on any subject dealing with their major field of study. Both sets of data were contributed by the same group of sophomores who were enrolled at Southwest Texas State Teachers College during the spring semester of 1951.

Methods Used in Comparing the Data

The sets of papers were compared primarily by correlation coefficients found by the application of the Pearson product-moment method. Probable errors of the correlation coefficients were determined by the use of the formula for the probable error of linear correlation (See Appendix).

All test scores were converted from the raw scores to T-scores as they are previously defined in this study. The conversion of the original composition ratings given to the compositions to a more convenient scale has been discussed in another section and, therefore, need not be treated here.

Results of Computations

All relationships were derived by computing correlation coefficients for various parts and combinations of parts of the two sets of original data. As a means to establishing validity, the marks which the students received in their freshman English courses, which dealt primarily with reading and writing, were also correlated with a part of the original data.

Composition Ratings as Criteria

The use of the composition ratings as criteria to determine whether the student's knowledge of written

composition is adequate was stated as a condition in the statement of the problem of this study. The question arose as to whether this assumption was advisable and well-founded. The best measure which was available to determine this validity was the marks received by these sophomore students in their freshman English work.

As a means to determining the validity of the composition ratings then, the ratings were correlated with the course marks. This correlation coefficient was found to be $.49 \pm .04$. Although not exceedingly high, the coefficient is significant; and it was judged advisable to use the composition ratings as criteria, but with discretion. The coefficient's value is raised somewhat if we consider (1) that one school year had elapsed between the time the course grades were given and the time that the compositions were written, (2) that since there is only one composition from each student this one effort may not coincide with the student's total writing ability, and (3) that although the freshman English courses teach primarily reading and writing, other factors are considered by the instructor in teaching the class and in assigning the marks.¹

The correlation coefficient for the scores on the standardized test, Total English, and the students' grades was found to be $.58 \pm .04$.

¹Paul B. Diederich, "The Measurement of Skill in Writing," School Review, LIV (December, 1946), 588.

As a means to establishing normality of this coefficient, it may be noted that Fletcher and Hildreth found a correlation of .48 between scores on the Ohio State University English Placement Test and instructors' ratings of students in English classes at Ohio State University.²

Relationships between the Two Sets of Data

Following this initial computation, the coefficients of correlation were computed for pairs of data as shown in Table III.

The three most significant coefficients were those of Effectiveness of Expression, Reading Comprehension, and Total English. The scattergrams given in Tables IV-X, which follow on pages thirty through thirty-six, show a more detailed picture of these relationships.

²Frank M. Fletcher and W. H. Hildreth, "Evaluation of an English Placement Test," Educational Research Bulletin, XXVIII (January, 1949), 7-17.

TABLE III
CORRELATION COEFFICIENTS FOR WRITTEN COMPOSITIONS
AND COOPERATIVE ENGLISH TESTS

X	Y	r	P.E. _r
Mechanics of Expression	Composition Ratings	.27	±.05
Effectiveness of Expression	Composition Ratings	.39	±.05
Reading Vocabulary	Composition Ratings	.30	±.05
Reading Speed	Composition Ratings	.27	±.05
Reading Comprehension	Composition Ratings	.36	±.05
Total Reading	Composition Ratings	.31	±.05
Total English	Composition Ratings	.38	±.05
Mechanics of Expression plus Effectiveness of Expression	Composition Ratings	.32	±.05

TABLE IV
EFFECTIVENESS OF EXPRESSION (ACTUAL NUMBERS)*

Composition Ratings	Number of Students in Divisions of 25%				Total
	0-24%	25%-49%	50%-74%	75%-99%	
6	9	14	18	24	65
5	10	10	16	11	47
4	1		1	1	3
3	11	10	4	2	27
2				1	1
1	7	3		1	11
0	1	2			3
Total	39	39	39	40	157

* Figures given in terms of actual number of students.

TABLE V
EFFECTIVENESS OF EXPRESSION (PERCENTAGES)*

Composition Ratings	Percentage of Students in Divisions of 25%				Total
	0-24%	25%-49%	50%-74%	75%-99%	
6	23.1	35.9	46.2	60.0	41.6
5	25.6	25.6	41.0	27.5	29.9
4	2.6		2.6	2.5	1.9
3	28.2	25.6	10.2	5.0	17.1
2				2.5	.6
1	17.9	7.7		2.5	7.0
0	2.6	5.2			1.9
Total	100.0	100.0	100.0	100.0	100.0

* Figures given in terms of percentage of students.

TABLE VI
 READING COMPREHENSION (ACTUAL NUMBERS)*

Composition Ratings	Number of Students in Divisions of 25%				Total
	0-24%	25%-49%	50%-74%	75%-99%	
6	9	14	16	26	65
5	14	11	15	7	47
4		2		1	3
3	9	7	8	3	27
2				1	1
1	4	5		2	11
0	3				3
Total	39	39	39	40	157

* Figures given in terms of actual number of students.

TABLE VII
 READING COMPREHENSION (PERCENTAGES)*

Composition Ratings	Percentage of Students in Divisions of 25%				Total
	0-24%	25%-49%	50%-74%	75%-99%	
6	23.1	35.9	41.0	65.0	41.6
5	35.9	28.2	38.5	17.5	29.9
4		5.2		2.5	1.9
3	23.1	17.9	20.5	7.5	17.1
2				2.5	.6
1	10.3	12.8		5.0	7.0
0	7.6				1.9
Total	100.0	100.0	100.0	100.0	100.0

*Figures given in terms of percentage of students.

TABLE VIII
TOTAL ENGLISH (ACTUAL NUMBERS)*

Composition Ratings	Number of Students in Divisions of 25%				Total
	0-24%	25%-49%	50%-74%	75%-99%	
6	6	17	16	26	65
5	11	13	16	7	47
4		1	1	1	3
3	14	4	5	4	27
2				1	1
1	7	2	1	1	11
0	1	2			3
Total	39	39	39	40	157

* Figures given in terms of actual number of students.

TABLE IX
TOTAL ENGLISH (PERCENTAGES)*

Composition Ratings	Percentage of Students in Divisions of 25%				Total
	0-24%	25%-49%	50%-74%	75%-99%	
6	15.4	43.6	41.0	65.0	41.6
5	28.3	33.3	41.0	17.5	29.9
4		2.6	2.6	2.5	1.9
3	35.9	10.3	12.8	10.0	17.1
2				2.5	.6
1	17.9	5.1	2.6	2.5	7.0
0	2.5	5.1			1.9
Total	100.0	100.0	100.0	100.0	100.0

* Figures given in terms of percentage of students.

CORRELATION OF COMPOSITION RATINGS AND EFFECTIVENESS OF EXPRESSION

36

Investigation of Causes

Part of the purpose of this study was to discover which section or sections of the standardized test, if any, could be most effectively used as a substitute for the composition-writing program. Since the difference between any two of the correlation coefficients is less than the sum of three probable errors (.15 in this case), the differences could have occurred by chance; and by this path of reasoning one might then assume that no one part has a closer relation to the composition ratings than does any other. Since the coefficients are relatively small, it might also be assumed that the two sets of data have little in common. However, since this is an exploratory investigation, all of the data being rather limited, the writer has assumed the data to be significant enough to warrant the investigation of possible causes and results of the differences which were found.

In Table III, page twenty-nine, it was shown that the most significant coefficient of correlation of individual test parts and the composition ratings was that of the test on Effectiveness of Expression which includes (1) sentence structure and style, (2) active vocabulary, and (3) organization, of which sentence structure and organization are listed specifically on the chart used in rating the compositions

originally. The least significant of the coefficients were those of Reading Speed (.27) and Mechanics of Expression (.27). Speed probably had little direct effect on the composition ratings since the students were given ample time--two hours--to write the compositions; but the Mechanics of Expression includes grammatical usage, punctuation and capitalization, and spelling, all of which are listed directly on the composition-rating chart. Perhaps, then, neither of the five characteristics discussed could be designated as the cause of the difference between the correlation coefficients since the parts from the tests with the lowest coefficients of correlation are specifically listed on the composition-rating chart as well as the parts from the test which yielded the highest correlation coefficient when compared with the composition ratings.

The only two parts remaining which were not directly listed on the composition-rating chart were active vocabulary and sentence style, which were both included in the test on Effectiveness of Expression. The question arose then as to the possibility of knowledge of vocabulary or sentence style being the chief cause for the difference. The possibility that knowledge of vocabulary might be the reason showed little validity when compared with the correlation coefficient for the Vocabulary test and the composition ratings which was .30. The ability to recognize and to use good sentence style was

the only other specific alternative. There were no specific scores relating only to sentence style, so the correlation of this characteristic and the composition ratings could not be determined. The writer only suggests this characteristic as a possible key to the cause for the difference in the coefficient significance, since it has not as yet been proved false.

The coefficient of correlation between the composition ratings and the test on Reading Comprehension (.36) is the next most significant factor, that is, excluding the Total English (.38) because it includes all parts of the standardized test and would not be specific. Reading Comprehension is not a single characteristic and cannot be classified as subject matter. Ability in reading comprehension rests on many factors; it is a complex function which requires several skills, abilities, and attitudes.³ From the correlation coefficient, it is noted that some of these skills, abilities, and attitudes are required for writing compositions. It seems also that these skills, abilities, and attitudes are some other than those measured by the tests on Mechanics of Expression or Effectiveness of Expression. Perhaps, also, the

³ Marjorie Seddon Johnson, "Factors in Reading Comprehension," Educational Administration and Supervision, XXXV (November, 1949), 385-406.

test on Effectiveness of Expression measures more than just the concrete skills listed in the contents; perhaps it measures some of the same factors which are taken into consideration in reading comprehension, since the correlation coefficient for the composition ratings and the test on Effectiveness of Expression is almost identical with the coefficient for the composition ratings and the test on Reading Comprehension.

The other differences found between the correlation coefficients are too insignificant to discuss except in a very limited manner, since they differ by not more than one probable error. They will not be dealt with here.

The Cutoff Point

One of the purposes of this study was to establish a cutoff point on the standardized test score-scale to mark the adequacy of the sophomore students' ability to do written composition. According to the composition-rating scale, the cutoff point which was used by the English department falls between ratings 2 and 3.⁴ This is shown in Table II, page

⁴Rating numbers enclosed in parentheses refer to original ratings given by the English department. Those not enclosed in parentheses refer to the converted rating derived for the purposes of this study.

eighteen. The department used the system whereby any student whose paper was given a single total grade of (3) in any rating combination was advised to do remedial work. Therefore, any student whose paper received rating combinations of (2), (2), and (3); (3) and (2); or (3) and (3) fell below this cutoff point. On the composition-rating scale, derived for more convenience in this study, this would include ratings of 0, 1, and 2.

The relationship of the standardized tests to this cutoff point can be noted by reference to Tables IV-X on pages thirty through thirty-six. The test on Effectiveness of Expression yielded the highest correlation coefficient when compared with the composition ratings, so its relationship would be the most significant. By cutting off papers for remedial work with composition ratings of 0, 1, or 2, as was done in the original program, it is noted that fifteen papers are included in this group; but three of those lie above the median on the test score-scale. There is no intention here to imply that the composition-rating results should be used as a criterion for judging the adequacy of the standardized objective test used. Indeed, the opposite may be true. Yet because of the low correlation of the two sets of data, it is probably not advisable to establish a cutoff point on either of the scales alone. If this were done, many students

who were rated high on composition would be advised to do remedial work because of their low scores on the standardized test.

It is suggested that a combination of the two sets of data be used in establishing a cutoff point. Table X, page thirty-six, gives a more complete picture of the correlation of the composition ratings and the test on Effectiveness of Expression. Approximately the low twenty-five per cent of the students on the composition-rating scale would lie within ratings of 0, 1, 2, and 3, and the low twenty-five per cent on the test score-scale would include raw scores of forty-two and below. If those students who fell in the low twenty-five per cent on both scales were required to do remedial work, the group would include approximately seventeen students or ten per cent of the entire 157 students. By using both sets of data, the validity of the composition-evaluating program is increased and the measure is more valuable.

CHAPTER III

SUMMARY AND CONCLUSIONS

It was the purpose of this study to investigate the possibility of substituting students' scores on a standardized English test for ratings on written composition as a means of evaluating the students' ability to do written composition. The investigation was made by comparing ratings given sophomores on written compositions with performance scores of these students on a standardized English test which was administered to them during the spring semester of 1951. These comparisons were made by correlation of the two sets of data, with a follow-up of observations by the writer.

All of the correlation coefficients of separate parts of the standardized test with the composition ratings fell between $.27 \pm .05$ and $.39 \pm .05$. These coefficients are not highly significant, and the writer concludes that it would not be advisable to substitute the standardized test scores for the composition ratings because the community factor between the two measuring techniques appears to be too small. However, the correlation coefficients are to some degree significant; hence, validity superior to either measuring technique would be obtained if the two were combined, thereby segregating students for remedial work who ranked in the low section of both sets of data.

APPENDIX

FORMULAE

1. Pearson Product-Moment Formula:

$$r = \frac{f_{xy}d_xd_y - \frac{(f_xd_x)(f_yd_y)}{N}}{\sqrt{\left[f_xd_x^2 - \frac{(f_xd_x)^2}{N} \right] \left[f_yd_y^2 - \frac{(f_yd_y)^2}{N} \right]}}$$

2. Probable Error for Linear Correlation Formula:

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

STANDARDIZED TEST SCORES* AND COMPOSITION RATINGS
(STUDENTS ARE INDICATED BY NUMBERS)

Student's Number	Mechanics of Expression (1)	Effectiveness of Expression (2)	Reading Vocabulary (3)	Reading Speed (4)	Reading Comprehension (5)	Total Reading (6)	Total English (7)	Composition Rating (8)
1	61	67	57	54	47	52	53	6
2	32	22	28	44	40	35	24	5
3	48	43	45	51	50	47	45	5
4	62	65	80	65	74	74	64	6
5	43	41	36	47	43	40	40	6
6	44	42	46	38	39	39	40	5
7	55	58	66	58	63	62	58	6
8	42	64	51	57	56	54	53	5
9	57	55	46	47	50	46	53	5
10	39	44	48	45	43	45	41	6
11	38	50	59	51	50	53	45	6
12	52	40	42	44	49	44	44	6
13	61	55	47	50	50	48	56	6
14	37	30	45	49	49	46	34	5
15	47	46	55	50	47	50	46	5
16	61	66	58	67	64	63	65	5
17	65	59	61	60	65	62	62	6
18	62	77	66	77	69	71	72	6
19	66	58	59	66	85	71	67	6
20	39	38	43	38	41	47	38	3
21	32	30	36	37	39	35	27	5
22	64	73	80	88	72	82	75	6
23	53	49	53	54	54	53	51	3
24	37	54	61	52	52	54	46	6
25	59	42	45	40	39	40	46	6

*All standardized test scores are in the form of T-scores as defined in this study.

Student's Number	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
26	53	67	58	54	54	54	54	6
27	49	32	48	39	43	41	38	5
28	43	56	52	41	46	45	46	5
29	58	49	61	49	52	53	53	5
30	53	62	55	51	54	53	57	5
31	63	59	53	68	70	64	54	6
32	44	46	40	41	45	40	42	3
33	65	58	59	71	58	63	62	6
34	53	55	58	50	49	51	53	5
35	64	77	80	78	69	77	75	6
36	65	65	72	71	68	71	69	6
37	50	67	49	45	44	45	51	5
38	71	63	66	60	58	62	67	6
39	38	41	49	51	52	50	41	3
40	53	50	46	44	45	44	49	4
41	38	54	45	44	49	45	44	5
42	55	45	38	44	44	40	46	6
43	47	47	47	57	53	52	48	3
44	64	52	55	57	60	57	58	6
45	48	38	42	39	40	39	38	5
46	60	53	55	59	56	57	57	6
47	55	53	41	49	52	46	57	6
48	59	49	65	56	50	56	54	3
49	44	37	34	38	37	34	36	1
50	37	38	45	40	43	50	38	6
51	48	45	48	65	54	56	49	5
52	53	65	68	55	65	63	67	6
53	54	67	52	48	49	48	53	5
54	53	44	42	35	32	34	43	5
55	36	33	38	35	37	34	30	3

Student's Number	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
56	48	46	48	52	56	51	48	6
57	33	34	40	34	35	34	29	0
58	40	49	65	50	68	60	48	6
59	63	61	55	69	63	63	64	6
60	51	42	48	47	44	45	45	5
61	63	58	51	50	56	51	58	5
62	48	45	46	42	43	42	44	5
63	59	67	55	54	56	54	57	5
64	42	50	73	50	65	63	50	1
65	58	45	53	49	45	48	51	6
66	47	49	51	47	49	47	46	5
67	51	45	48	44	47	45	46	3
68	52	63	51	69	68	63	60	6
69	62	51	54	59	58	57	56	5
70	57	54	59	55	50	54	56	5
71	52	50	49	48	50	48	50	3
72	51	44	48	42	44	44	45	6
73	47	41	42	39	37	38	41	3
74	51	61	53	49	52	51	54	5
75	44	41	48	51	50	48	43	6
76	43	62	51	55	52	52	52	5
77	43	42	46	35	34	35	38	3
78	61	52	59	62	47	56	57	6
79	59	56	52	49	47	48	54	5
80	48	56	72	60	68	66	56	6
81	51	59	73	62	65	66	59	6
82	47	45	61	48	50	52	46	6
83	42	46	48	45	43	45	43	0
84	51	56	53	57	51	53	53	6
85	41	33	40	45	47	44	36	6
86	64	62	48	59	54	53	61	6
87	58	67	57	62	52	57	36	6
88	37	39	40	42	47	48	36	3
89	37	50	39	44	37	38	38	3
90	54	42	51	44	44	45	46	6

Student's Number	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
91	53	62	59	75	65	66	61	6
92	40	46	55	44	49	48	43	6
93	69	55	68	58	58	62	64	4
94	39	41	48	55	45	48	41	1
95	37	44	49	44	50	46	41	3
96	43	42	51	50	44	47	42	6
97	59	58	59	55	45	52	57	5
98	37	41	43	44	43	42	37	5
99	70	60	54	54	52	53	62	6
100	57	59	71	79	62	71	62	3
101	54	53	51	44	50	47	51	6
102	33	40	24	36	35	28	29	6
103	41	50	59	40	44	46	44	5
104	59	45	46	49	45	46	50	5
105	58	51	48	49	44	46	52	6
106	39	47	49	41	43	44	42	1
107	59	54	47	41	44	42	52	6
108	65	52	48	64	44	51	57	3
109	54	49	51	50	44	47	50	6
110	55	67	58	57	56	57	57	5
111	38	38	40	42	47	41	36	6
112	54	59	61	55	47	53	57	3
113	37	33	38	37	40	35	32	3
114	47	40	53	41	44	45	42	4
115	61	54	55	50	50	51	56	3
116	47	43	47	45	44	44	43	6
117	32	34	42	28	26	28	30	3
118	34	36	38	48	44	41	34	1
119	51	49	46	64	51	53	51	5
120	46	35	41	36	26	32	35	1
121	41	44	39	41	43	39	40	5
122	39	37	36	45	46	41	36	3
123	63	61	53	55	65	57	61	5
124	37	50	42	41	39	39	38	3
125	67	63	59	67	63	63	66	6

Student's Number	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
126	42	51	42	44	44	41	43	5
127	38	34	40	39	26	32	30	1
128	62	61	61	59	58	59	65	6
129	54	68	58	65	53	59	61	6
130	38	56	34	47	54	44	52	5
131	48	65	61	66	68	65	60	6
132	62	62	53	55	56	54	60	5
133	44	32	45	49	47	46	38	1
134	43	44	46	44	43	42	42	6
135	44	43	43	44	44	42	42	3
136	43	39	38	45	44	41	38	1
137	49	47	47	55	56	52	49	3
138	67	60	53	45	50	48	60	6
139	64	66	61	50	58	56	64	3
140	52	47	48	54	49	50	51	6
141	50	50	47	50	58	51	50	6
142	48	47	48	41	35	40	44	0
143	58	58	46	49	49	47	54	5
144	39	31	43	39	43	40	34	6
145	35	38	42	42	44	41	35	3
146	62	54	46	45	50	46	54	6
147	48	51	52	42	43	45	48	5
148	63	68	70	56	74	70	69	5
149	44	38	38	44	41	39	38	5
150	39	44	42	42	44	41	40	1
151	42	54	38	44	44	40	44	6
152	61	62	66	75	59	66	64	2
153	38	41	45	41	43	41	37	5
154	49	47	55	41	45	46	46	5
155	61	56	73	44	49	54	58	6
156	62	62	61	64	57	60	62	1
157	39	42	64	45	54	53	43	3

BIBLIOGRAPHY

Books

- Bachelor, Joseph M., and Ralph L. Henry, American Thinking and Writing, D. Appleton-Century Company, New York, 1942.
- Canby, Henry Seidel, English Composition in Theory and in Practice, The Macmillan Company, New York, 1920.
- _____, and John Baker Opdycke, Elements of Composition for Secondary Schools, The Macmillan Company, New York, 1920.
- Mirriclees, Lucia B., Teaching Composition and Literature, Harcourt, Brace and Company, Chicago, 1937.
- Pence, Raymond W., College Composition, The Macmillan Company, New York, 1929.
- Rankin, Thomas E., Clarence D. Thorpe, and Melvin T. Solve, College Composition, Harper and Brothers, New York, 1929.
- Reed, Percy Isaac, and Elizabeth Frost Reed, Applied Composition, Thomas Nelson and Sons, New York, 1936.
- Shaw, Harry, Complete Course in Freshman English, The Macmillan Company, New York, 1929.
- Sheridan, Bernard M., Speaking and Writing English, Benjamin H. Sanborn Company, Boston, 1927.
- Shipherd, Robinson H., The Fine Art of Writing for Those Who Teach It, The Macmillan Company, New York, 1926.
- Thomas, Charles Swain, The Teaching of English, Houghton Mifflin Company, New York, 1917.

Periodical Articles

- App, Austin J., "Old Ben Jonson on Grading Compositions," College English, XIV (February, 1943).

- Diederich, Paul B., "The Measurement of Skill in Writing," School Review, LIV (December, 1946).
- Edmiston, R. W., and C. N. Gingerich, "The Relation of Factors of English Usage to Composition," Journal of Educational Research, XXXVI (December, 1942).
- Fletcher, Frank M., and W. H. Hildreth, "Evaluation of an English Placement Test," Educational Research Bulletin, XXVIII (January, 1949).
- Henry, George, "How to Get Interesting Themes," The English Journal, XXIII (September, 1944).
- Johnson, Marjorie Seddon, "Factors in Reading Comprehension," Educational Administration and Supervision, XXV (November, 1949).
- Turner, Lucile, "Theories and Practices in Evaluation and Correction of Oral and Written Composition," Peabody Journal of Education, XIX (March, 1942).
- Wykoff, George S., "Suggestions for the Reading of Themes," College English, XXI (January, 1950).