INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/4017

76-3097

FULLER, Sarah L., 1921-
AN ANALYSIS OF SELF-EVALUATIONS IN DETERMINING SECONDARY SCHOOL STUDENTS' GRADES.

The University of Oklahoma, Ed.D., 1975
Education, administration

Xerox University Microfilms, Ann Arbor, Michigan 48106

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED.

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

AN ANALYSIS OF SELF-EVALUATIONS IN DETERMINING SECONDARY SCHOOL STUDENTS' GRADES

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF EDUCATION

BY

SARAH FULLER

Norman, Oklahoma

1975

AN ANALYSIS OF SELF-EVALUATIONS IN DETERMINING SECONDARY SCHOOL STUDENTS' GRADES

APPROVED BY

DISSERTATION COMMITTEE

TABLE OF CONTENTS

LIST OF TABLES vi

LIST OF FIGURES vii

LIST OF APPENDICES viii

Chapter

I. INTRODUCTION AND STATEMENT OF THE PROBLEM . . 1

Introduction 1
Statement of the Problem 6
Hypotheses Tested in the Study 6
Definition of Terms 8
Limitations of the Study 9
Statistical Design of the Study 10
Procedure of the Study 12
Organization of the Study 12

II. REVIEW OF RELATED LITERATURE 14

Early Attempts at Self-Evaluations 15
Recent Literature on Student Self-Evaluation 17
Feedback and Self-Evaluation 30
Practicing Student Self-Appraisal 36
The Student Self-Report Form 37
Student Involvement in Evaluation 39
Class Evaluations of Compositions 43
The Conservative Approach to Student Self-Evaluations 45

III. METHODS AND PROCEDURES 49

Pre-Experimental Procedures 50
Choice of Research Design 50
Development of a Self-Evaluation Rating Sheet 51
Establishing the Reliability and Validity of the Data Collection Instrument 54

TABLE OF CONTENTS (Cont'd)

		Page
	Training the Students to Conduct Self-Evaluations of Their Progress	55
	Selection of the Participants	56
	Experimental Procedures	57
	Data-Analysis Procedures	61
	Choice of Statistical Procedures	61
IV.	RESULTS OF STATISTICAL ANALYSIS	64
	Descriptive Statistics of the Four Classes	65
	Results of Testing the Hypotheses	65
	Results of Testing the First Null Hypothesis	65
	Results of Testing the Second Null Hypothesis	67
	Results of Testing the Third Null Hypothesis	69
	Results of Testing the Fourth Null Hypothesis	71
	Results of Testing the Fifth Null Hypothesis	74
	Results of Testing the Sixth Null Hypothesis	75
V.	SUMMARY, CONCLUSIONS, AND IMPLICATIONS FOR FURTHER RESEARCH	79
	Summary	79
	Conclusions	81
	Implications for Further Research	83
BIBLIOGRAPHY		85

LIST OF TABLES

Table		Page
1	Descriptive Statistics of Grades Estimated by Students, Grades Assigned by the Course Instructor, and the Arithmetic Differences . . .	66
2	A Comparison of Course Grade Estimates Made by the Experimental Groups and the Actual Course Grades Assigned by the Instructor	68
3	A Comparison of Course Grade Estimates Made by the Control Groups and the Actual Course Grades Assigned by the Instructor	70
4	A Comparison of the Estimated/Actual Course Grade Differences Computed for the Experimental and Control Groups	70
5	The Relationship Between the Course Grade Estimates Made by the Experimental Groups and the Actual Course Grade Assigned by the Instructor	73
6	The Relationship Between the Course Grade Estimates made by the Control Groups and the Actual Course Grades Assigned by the Instructor	76
7	A Comparison of the Correlation Coefficients Computed Between Estimated and Actual Grade Scores for Experimental and Control Groups . . .	78
8	Raw Data and Descriptive Statistics of Student Group Number One (Experimental Group One) . . .	94
9	Raw Data and Descriptive Statistics of Student Group Number Two (Experimental Group Two) . . .	95
10	Raw Data and Descriptive Statistics of Student Group Number Three (Control Group One)	96
11	Raw Data and Descriptive Statistics of Student Group Number Four (Control Group Two)	97

LIST OF FIGURES

Figure		Page
1	Research Design Used in the Experiment	52
2	Sampling Paradigm of the Experiment	58
3	Inferential Statistics Chosen to Test Each Null Hypothesis	63

LIST OF APPENDICES

Appendix		Page
A	General Objectives Adopted for the Five-Unit Study Project	88
B	The Student's Self-Evaluation Rating Sheet Used by the Students Who were Taught Self-Evaluation Techniques	90
C	Raw Scores and Descriptive Statistics of the Four Student Groups	92

CHAPTER I

INTRODUCTION AND STATEMENT OF THE PROBLEM

One of the most controversial practices in public schools today is the grading system used in reporting students' progress. Assigning letter grades or comparable indicators of each student's academic performance is usually the classroom teacher's responsibility. After teachers have issued the student grades, they very often receive criticism. This criticism, generally from the students, students' parents, and administrators, may cause the teacher to approach the next grade reports with dread and trepidation. Yet, the conscientious teacher usually gives each student's grade careful consideration. Many factors such as the student's ability, general attitude, and motivation may be considered before a final grade is assigned. The following quotation from a study by Barnes and Barnes is an example of the forethought which goes into assigning student grades:

> I know he should get an A but I just can't see my way clear to give it to him. I know she didn't quite come up to a C, but she tried so hard I felt I had to give it to her.[1]

Such teacher comments are indicative of the confusion and subjectivity associated with students' grades. Grades

---

[1]Katherine F. Barnes and E. H. Barnes, "A Realistic Approach to Grading," The Clearing House, (April, 1962), XXXVI, p. 476.

usually reflect the frustrations of a teacher who is trying to be fair, who is trying to make an objective appraisal of the student's academic achievement, who is considering the individual effort made by the student, and who is considering the personal qualities of the student. While all these factors should be taken into consideration, the resulting student grades can be very subjective and unfair. For instance, Crowley found significant inconsistencies among the marking (grading) practices of teachers in Oklahoma City's Public Secondary Schools. He concluded that; "The grades pupils receive appear to lose their meaning because of a lack of clearly defined marking practices and procedures."[2] The absence of clearly defined grading practices may lead to further complications. For example, Grambs, Carr and Fitch found significant differences among the values different teachers placed on letter grades. They concluded that "...research and experience show that an A to one teacher may be a C to another."[3] Barnes and Barnes also recognized this inconsistency and advocated uniformity in grading practices. One conclusion drawn from their study was as follows:

---

[2]Leon V. Crowley, "An Investigation of the Marking Practices in Secondary Schools," (unpublished Doctoral Dissertation, College of Education, University of Oklahoma, Norman, Oklahoma), 1968.

[3]Jean D. Grambs, John C. Carr, and Robert M. Fitch, Modern Methods in Secondary Education, (New York: Holt, Rinehart, and Winston, Inc., 1970), p. 328.

> If teachers in a given school exercise a uniformity in their grading practices, it matters little whether a given performance is graded A or B, because the relative rank of the graduates will not be affected.[4]

Such inconsistencies sometimes occur because teachers use different criteria in assigning students grades. Practices often consist of amassing scores from such areas as daily assignments, tests, essays, reports, term papers, and term projects. Grades from some of these sources may be given more importance (weight) than the grades from other sources depending on the value the teacher places on a particular assignment. Averaging the final scores to determine the student's grade may result in a distorted view of his total performance. Such procedures may produce grades which can only be interpreted by the classroom teacher who assigned them. The following statement made by Wrinkle seems to substantiate this:

> So many different factors entering into the determination of a student's mark are considered as of such varying importance by different teachers that it is almost impossible for anyone except the teacher who gave the mark to tell what it means.[5]

Since the teacher who assigns certain grades to student performances is the only one who can interpret them, they are often called upon to perform such an interpretation.

---

[4] Barnes and Barnes, p. 479.

[5] William L. Wrinkle, Improving Marking and Reporting Practices, (New York: Holt, Rinehart and Co. Inc., 1947), p.11.

The subjectivity involved in the present grading systems has drawn the attention and criticism of many individuals and special interest groups who have contended that grading may be a violation of the students' civil rights.[6] Through such efforts, students have become increasingly aware of their civil rights. While most violations of students' rights are associated with other areas, it is not beyond the realm of possibility that some grading systems can and do violate students' rights. In light of such possibilities, there have been some attempts to assist those students who feel that the school system has violated their civil rights.[7] These developments are a strong indication that it is simply a matter of time before students will have an active role in determining the type of evaluation/grading system used by a particular school. Moreover, the students' role will probably be to assist in determining their own grades.

This poses another problem, however. Few students have sufficient training to participate in the student grading evaluation process. While there have been attempts to develop grading schemes which would allow student input, none has been widely accepted.

Sawin conducted experimental studies with student

---

[6]Gerald Melson. Aspects of the Legal Relationship Between Students and the Public High School. (Norman: The University of Oklahoma Press, 1970), p. 29.

[7]Ibid., p. 28.

self-evaluation techniques. However, Sawin's efforts were directed toward developing a system for training students and teachers in self-evaluation techniques and not in the development and refinement of self-evaluation procedures.[8]

Grambs, Carr, and Fitch also conducted research studies in the area of student self-evaluation. Their method could more properly be regarded as a small-group/self-evaluation technique. However, Grambs et al. were concerned with student self-evaluations of daily assignments and did not consider the students' ability to evaluate their overall class performance.[9]

The studies by Sawin and Grambs et al. did not consider the students' ability to evaluate their class performance over a prolonged period of time. While the results of these studies indicated that students can evaluate their performance on daily assignments, it cannot be safely assumed that they could do equally well when evaluating their performance over a longer period of time.

The limitations of such previous research studies suggest many questions. How proficient are high school students at evaluating their grades in a particular course? Can they be taught self-evaluation techniques which would allow a systematic assessment of class performance? How accurately can high

---

[8]E. I. Sawin, Evaluation and the Work of the Teacher. (Belmont, California: Wadsworth Publishing Co., Inc., 1969). p. 193.

[9]Grambs, Carr, and Fitch, op. cit., p. 411.

school students estimate the grades they will receive in a particular course?

Statement of the Problem

The problem of this study was to conduct an analysis of self-evaluations in determining secondary school students' grades. The study compared the differences between student self-evaluations and teacher's grades in classes utilizing self-evaluation techniques with differences between self-evaluations and teacher's grades in classes not utilizing self-evaluation techniques.

Hypotheses Tested in the Study

In order to achieve the purposes of the present study, several hypotheses were tested for significance at the .05 level. In order to determine the benefits of teaching students self-evaluation techniques, three hypotheses were stated. The first hypothesis was to determine how accurately the students who had been taught self-evaluation techniques could estimate their course grades; the second was to determine how accurately the students who had not been taught self-evaluation techniques could estimate their course grades; and the third was to determine whether there was a significant difference between the two groups' errors in grade estimation. The first three null hypotheses were stated as follows:

Ho_1 There is no statistically significant difference between the course grade estimates made by high school students who are taught self-evaluation techniques and the actual

course grades assigned to them by their instructor.

Ho_2 There is no statistically significant difference between the course grade estimates made by high school students who are not taught self-evaluation techniques and the actual course grades assigned to them by their instructor.

Ho_3 There is no statistically significant difference between the estimated/actual course grade differences of high school students who are taught self-evaluation techniques and the estimated/actual course grade differences of high school students who are not taught self-evaluation techniques.

Three additional hypotheses were tested in the study. However, these were tested as an attempt to determine the correlation between the two student groups' actual and estimated course grades. Whereas the first three hypotheses were tested to determine the amount of difference between their actual and estimated course grades.

The fourth null hypothesis was to determine the correlation between self-evaluating students' grade estimates and actual course grades; the fifth hypothesis was to determine the correlation between the estimated and actual course grades of students who had not been taught self-evaluation techniques, and the sixth hypothesis was to compare the correlation coefficients computed for the two student groups. The final three null hypotheses were stated as follows:

Ho_4 There is no statistically significant relationships between the course grade estimates made by high school students who are taught self-evaluation techniques and the course grades assigned to them by their instructor.

Ho_5 There is no statistically significant relationship between the course grade estimates made by high school students who are not taught self-evaluation techniques and the course grades assigned to them by their instructor.

Ho_6 There is no statistically significant difference between the correlation coefficient computed between the estimated and actual course grades of the two experimental groups (those students who were taught self-evaluation techniques) AND the correlation computed between the estimated and actual course grades of the two control groups (those students who were not taught self-evaluation techniques).

Definition of Terms

In order to avoid confusion and multiple interpretations of some of the terms used in the present study, the following definitions and explanations were established.

(1) High School Student: The 102 students enrolled in American Literature at Southeast High School in Oklahoma City, Oklahoma for the Fall semester of the 1972-73 academic year.

(2) Course Instructor/Teacher: The instructor teacher employed by the Oklahoma City School System to teach the American Literature class for the Fall semester of the 1972-73 academic year.

(3) High School English Course: American Literature course offered by Southeast High School in Oklahoma City, Oklahoma during the Fall semester of the 1972-73 academic year.

(4) Actual Course Grade: The alphameric mark assigned by the course instructor for a student's class performance in the high school English course.

(5) Estimated Course Grade: The alphameric estimate

made by the Experimental and Control Groups of the grade they would receive from their performance in the high school English course.

(6) Estimated/Actual Course Grade Difference: The arithmetic difference between the students' Estimated Course Grades and their Actual Course Grades.

(7) Self-Evaluation Techniques: The techniques taught to the high school students whereby they could make a systematic assessment of their performance in the high school American Literature class by using the Self-Evaluation Rating Sheet shown in Appendix B.

(8) Experimental Groups/Classes: The two groups (classes) of high school students who were taught self-evaluation techniques.

(9) Control Groups/Classes: The two groups (classes) of high school students who were not taught self-evaluation techniques.

(10) Self-Evaluation Rating Instrument: The evaluation instrument shown in Appendix B.

Limitations of the Study

In order to make the present study possible, it was necessary to establish several parameters or limitations. These limitations were used to define the study participants, limit the statistical generalizations to the present student groups, and specify the size, location, and time of the entire study.

The most important limitations placed on the study were as follows:

(1) The population of students was limited to four (4) class groups (N=102) who were enrolled in American Literature #102 at Southeast High School in Oklahoma City, Oklahoma for the Fall semester of the 1972-73 academic year.

(2) The course grades estimated by the student population and assigned by the course instructor were limited by the grading system which utilizes the letter grades of A, B, C, D, and F. Alphameric grade estimates were assigned the following weights: A = 4.0, B = 3.0, C = 2.0, D = 1.0, and F = 0.0. These discrete number categories would not allow the investigator to reflect partial grades such as a B+ or a C-.

Statistical Design of the Study

To ascertain the differences between the grades selected by the self-evaluating students and the teacher's marks for them and the grades selected by the non-self-evaluating students and the teacher's marks for them, a descriptive survey and analytical study was made. The data were coded to facilitate analysis. Each datum collected was measured for strength of correlation with the Pearson product-moment correlation coefficient. In this way it was possible to compute the amount of relationship between the

students' self-evaluations and teacher's grades for them. A contingency coefficient was computed between the estimated scores of non-self-evaluating students and teacher's grades for them.

Weinberg and Schumacker, recommend this type of statistical treatment.

> The correlation coefficient (often called the Pearson product-moment correlation coefficient), signified by the letter "r" is a precise measure of the way in which two variables correlate. Its value is such as to indicate both the direction (postive or negative) and the strength of the correlation between two variables.[10]

Kerlinger concurs that the correlation coefficient is appropriate for this type of statistical analysis.

> Product-moment and related coefficients of correlation, then, are based on the concomitant variation of the members of sets of ordered pairs. If they covary, vary together--high values with high values, medium values with medium values, and low values with low values, or high values with low values, and so on--it is said that there is a positive or negative relation as the case may be. If they do not covary, it is said there is "no" relation.[11]

To perfect the statistical analysis of this study, the investigator enlisted the aid of a statistician/computer analyst to aid in writing a computer program. Using a computer helped to assure the accuracy of the results.

---

[10]George H. Winberg and John A. Schumacker, Statistics, An Intuitive Approach, (Belmont, California: Wadsworth Publishing Company, Inc., 1967), p. 259.

[11]Fred N. Kerlinger, Foundations of Behavioral Research (new York: Holt, Rinehart and Winston, Inc., 1967), p. 90.

Procedure of the Study

The procedure for the study utilized the methodology of action or operational research. Good defines action research as "an attempt to provide investigational procedures suitable for study and solution of school problems in relation to the total situation."[12] The data were collected through the use of a questionnaire. Wiersma defines questionnaire as "a list of questions or statements to which the S is asked to respond by a written response."[13] He compares the questionnaire to the personal interview.

> A questionnaire is sometimes referred to as a written, self-administered interview, and by the same token we could consider an interview as an oral questionnaire. The two types of data collection have a great deal in common relative to item construction and use.[14]

Organization of the Study

The background and rationale for the study were discussed earlier, yet there is a need to deal with the background of student self-evaluation practices in greater detail. A second division of the study presents details in a review of the literature. The third chapter is devoted to the development of the Student's Self-Evaluating Rating Sheet instrument, methods, and procedures used in the study.

---

[12] Carter V. Good, Essentials of Educational Research, Methodology and Design, (New York: Appleton-Century-Crofts, 1966), p. 262.

[13] William Wiersma, Research Methods in Education, An Introduction, (New York: J. B. Lippincott Company, 1969), p. 274.

[14] Ibid., p. 274.

Chapter IV deals with the data collected and the statistical treatment of those data. Chapter V reports the findings of the study, conclusions reached, and recommendations for the development of a valid approach to student self-evaluation.

CHAPTER II

REVIEW OF RELATED LITERATURE

This study was concerned with eleventh-grade high school students' ability to evaluate their performance in English literature classes of a large metropolitan high school setting. A survey of the literature was made and a collective synthesis of the related studies is presented in this Chapter of the dissertation.

DiSibio believed he had the correct formula for student evaluation when he wrote the following:

> Today's perennial topic of debate: the grading system lies in the hands of a new idea in grading; student evaluation plus teacher scrutinization plus teacher-parent-student conferences. The utilization of this formula could be what education is looking for and needs.[1]

In reading DiSibio's statement, careful attention should be given to the terminology used. For instance, he referred to the present grading system as "Today's perennial topic of debate:" DiSibio also used the term "evaluation" for students and "scrutinization" when referring to teachers. Criticism of the marking process is indeed perpetual or never-ending, and this survey of related literature has followed the

---

[1]Robert A. DiSibio, "A Different Approach to Grading," Education, Vol. 92, 2 (Nov.-Dec., 1972), p. 35.

development of the idea of student self-evaluation, since it seemed to offer a more democratic method of grading than having the teacher assign grades.

Early Attempts at Self-Evaluation

In early attempts to get students to assess their own performances in an academic subject, Symonds stated, "The suggestion that pupils evaluate themselves is somewhat new and is attracting considerable attention."[2]

Symonds cited the 1931 study by Shaw which showed that "students estimated their achievement with considerable error."[3] Asch allowed students to request their own grades and found the majority honest. Some, he added, "requested A or B but had very limited knowledge of the subject matter."[4] In his book Faunce commented in a section on evaluation, "At the present time, the evaluation program relative to core classes in high school is carried out by the teacher."[5] Inlow conducted a comparative study in which he described student teachers evaluating themselves and then compared their ratings with critic teachers and university

---

[2]Percival M. Symonds, "Pupil Evaluation and Self-Evaluation," Teachers College Record, LIV, (December 1952), p. 141.

[3]Ibid.

[4]Ibid.

[5]Roland C. Faunce and Nelson L. Bossing, Developing the Core Curriculum (New York: Prentice-Hall, 1951), p. 272.

supervisors. He described the student teachers' ratings as "unrealistic."

> The problem of the study was to determine to what extent the evaluations of student-teaching outcomes are in agreement when student teachers, critic teachers, and university supervisors are the three rating groups. The instrument which was used for the rating was an eight-item questionnaire with a five point continuum scale provided for the answers. The sample consisted of 45 student teachers of Northwestern University who spent a half day for one semester in one of the cooperating schools or two hours a day for two quarters.
>
> The results indicated that student teachers are unrealistic as self-raters inasmuch as their rating correlated only slightly with those of the critic teachers and the university supervisors. The coefficients for the ratings of the critic teachers and the university supervisors were substantial, but not high. The rating instrument gave indications of being reasonably effective if used by professional observers.[6]

Despite the findings of the study, Inlow was not discouraged concerning the value of self-evaluations and described improved techniques in a book published in 1970. These techniques are reviewed later in the Chapter.

In 1953, Russell reviewed the literature on self-evaluation.

> Some ten years ago, in a summary of existing practices, Hamalainen concluded, "The extent to which a pupil should contribute to his own evaluation is not clear. It is certain that he should enter into the process; yet how and to what extent is only

---

[6]Gail Inlow, "Evaluating Student Teacher Experience," Journal of Educational Research, Vol. XLV, 9 (May, 1952), p. 713.

partially understood at the present.[7]

This reference inferred that self-evaluation was seen as a useful technique as early as 1943. The self-evaluation practices, however, of the 1950's were primarily theory with little practical testing. Russell confirmed this idea.

> Although self-evaluation has been commonly accepted at the verbal level, it seems to be rare at the action level. The comment of Orata illustrates this lag between theory and practice: "Evaluation, as an integral part of the learning process, is much more talked about than practiced." This lack of emphasis upon self-evaluation in the school's appraisal program may be due to indifference, lack of knowledge, or difficulties in using a procedure unsuited to elementary and secondary school pupils and to current school practices.[8]

In the literature educators urged teachers to allow students to evaluate themselves, but generally teachers did not or would not follow the suggestions. Russell reviewed the Bullock study of 1948 which allowed students in grades 4-8 to evaluate themselves in reading, spelling and handwriting. He found "students generally rated themselves higher than their teachers."[9] Russell found no research of self-evaluations of senior high school pupils. He did state, however, that further investigation was in order.

---

[7]David H. Russell, "What Does Research Say About Self-Evaluation?" Journal of Educational Research, XLVI, 8 (April, 1953), p. 561.

[8]Ibid, p. 562.

[9]Ibid, p. 570.

> The glib statement that self-evaluation is an important part of all evaluation programs would seem to need both further investigation and extreme caution in its application.[10]

Grambs gave extensive treatment in the literature to self-evaluations as early as 1951. The "Am I Growing" report form, often used in student self-evaluations, was in print as early as 1941. In using the self-evaluation chart, the student began to assume the responsibility for his own improvement and growth in classroom practices. Group evaluation techniques also were recommended by Grambs.

> An evaluation committee with constantly rotating membership may be a great asset to a teacher. The students then act in an advisory capacity in judging the quality of their own work. For example, it has been found very rewarding for small groups of students - three or four - to read and evaluate each other's papers.[11]

Grambs believed that student involvement in self-evaluations developed the "ability to appraise and discriminate and set their standards and goals for achievement."[12] She believed that the greater the part the student played in evaluating himself, the more effective his learning would be.

Alexander considered the role of teacher-student conferences and how they help students to evaluate themselves

---

[10] Ibid.

[11] Jean D. Grambs and William J. Iverson, Modern Methods in Secondary Education (New York: William Sloane Associates, 1952), p. 389.

[12] Ibid., p. 390.

honestly and accurately. Of major importance in developing self-evaluation techniques in the student is the teacher-counseling conference. Alexander preferred indirect methods of teacher counseling to director formal counseling.

> Indirect methods of teacher counseling can be of very great value to youth in thinking through their own problems, goals, and accomplishments. Teachers who make effective use of this aid to self-evaluation become adroit in conferencing with pupils to push the latter's thinking on such successive questions as, What improvements are you trying to make in the class? Why are you trying to make these improvements? How are you going about it? Do you think you are making progress?[13]

Alexander cautioned that these conferences must be conducted without the tension of an impending grade or without a listening audience. The rapport between teacher and student must be relaxed and sincere with the student's discussion free and open.

Multiple techniques in student self-evaluations were offered by Alexander. He suggested self-marking charts and check lists which may be used over several weeks' time or question-answer forms for daily use. His "Individual Pupil Evaluation" for social studies is an adaptable self-marking progress chart that may be utilized for any subject. The instrument used in this study drew heavily from this valuable form. Of course, the items evaluated in English were quite different from those Alexander used in social studies, but

---

[13]William M. Alexander and Paul M. Halverson, Effective Teaching in Secondary Schools (New York: Holt, Rinehart and Winston, Inc., 1957), p. 402.

the rating process was similar. His "How Did I Do Today?" list of questions was applicable to daily progress of the student. This list contained the following questions:

> Did I get said what I wanted to?
> Did I talk more than I should have?
> Did I talk less than I should have?
> Were the ideas I presented worth presenting?
> Did I help the class make progress?
> Did I pay attention to others?
> Did I help to make this class worthwhile?[14]

Alexander also suggested questions that students may use from time to time to induce some self-analysis in an actual class in local government. These questions were as follows:

> What are the chief ideas about government in _____ that you have required as a result of this unit?
>
> Explain how these ideas have helped or will help you.
>
> What other questions do you have about government in _____?
>
> How can you find out the answers to these questions?
>
> Are you satisfied with your work in this unit? Why or why not?[15]

These questions were adaptable and could be varied depending on the course and the students. They allowed the student a chance to deviate from a chart or check list and elaborate concerning his progress or lack of progress.

---

[14] Ibid., p. 404.

[15] Ibid., p. 405.

Recent Literature on Student Self-Evaluations

Cautious about the worth of student evaluations, Ahmann and Glock concluded that students need to be carefully trained in making sound judgments regarding self-evaluations.

> . . . teachers should guide students in learning how to evaluate themselves. This is an important educational objective, and one helpful way of achieving it is to discuss a pupil's progress with him. He should be able to interpret his achievement as it relates to his strengths and weaknesses.[16]

By 1971, Ahmann and Glock conceded that it was necessary for students to gain experience in self-evaluations if they were to develop the ability to evaluate themselves impartially.

> Pupils as well as teachers can judge performances---their own and also those of their classmates. Much can be said for this type of multiple evaluation as a learning experience. When both pupils and teachers use the same instruments to judge the same performance simultaneously, and the results are examined, the pupils gain a much better perspective of the important features of the performance and the teacher's standards concerning them.[17]

Not all recent findings, however, have been restricted to multiple evaluations. Ahmann and Glock reflected the educator's lack of trust in the individual student's ability or integrity in judging himself. Wilhelms, however, believed

---

[16]Stanley Ahmann and Marvin D. Glock, Evaluating Pupil Growth (Boston: Allyn and Bacon, Inc., 1963), p. 579.

[17]J. Stanley Ahmann and Marvin D. Glock, Evaluating Pupil Growth (Boston: Allyn and Bacon, Inc., 1971), p. 235.

this lack of trust lowered the student's confidence in himself.

> Whether it is organizational, instructional or evaluative, any practice which limits the opportunity or lowers the confidence of the learner in evaluating his own progress and purposes is detrimental to his development and training.[18]

The student needs experience in evaluating himself to gain confidence in his self-evaluation abilities. When that confidence comes, the student has made a major step forward in assuming responsibility for his own improvement. The opportunity must be provided for the learner to gain experience in evaluating himself, Wilhelms concluded.

> We must respect and believe in the learner's ability to evaluate himself and his learning and to improve his evaluation if he has reasonable support and assistance. The longer the learner has experienced only outside evaluation the harder it will be to get him to take the initiative.[19]

Wilhalms also emphasized that the self-evaluation must be continuous to give the student a realistic basis for his self-directed activity and for his involvement in his own learning.

Part of the technique of self-evaluation includes individual conferences to provide opportunities for the teacher and the learner to evaluate together, Wilhelms stated. The teacher was urged to look at the learner's goals and specific purposes, to decide where the student was in the learning

---

[18]Fred T. Wilhelms, ed., Evaluation as Feedback and Guide (Washington, D. C., 1971), p. 96.

[19]Ibid., p. 98.

process, what has been mastered, what was in process, and what the next steps for the learner might be. Interaction between teacher and student can be a fruitful endeavor if the student views his role and his progress positively. Evaluative interaction can be a positive force in personal development if the following conditions are met:

1. When that evaluation is in terms of what is important to the learner.
2. When it is a means which he can use for directing his future study.
3. When it gives him feedback to help him know his competencies, where he is, to what point he has progressed.
4. When he feels that the teacher is helping him evaluate himself and improve his perceptions as to his next needs.
5. When it helps him see next steps and opens doors for him to move forward; not when it stops study and closes doors to future learning in the area.
6. When it helps him gain personal meaning in an area of learning.
7. When it helps him see progress and thus feel good about himself, even if he realizes he still has some way to go.
8. When it encourages rather than discourages further learning.
9. When it is set up so as to be a challenge leading to what he believes he can do, rather than a frustration convincing him he cannot succeed.
10. When in one form or another it is continuous.
11. When the learner comes to see the behavior valued by the teacher to be his own realistic self-appraisal and self-direction, rather than convincing the teacher of the amount of

information learned.[20]

Although Clark and Starr did not deal with student self-evaluation as extensively as Wilhelms, they agreed that students' goals set in a course of study and teacher-student conferences were advantageous in training for self-evaluation. Clark and Starr advised students to observe certain aids to self-evaluation.

1. Keep anecdotal reports to measure work.
    a. Keep a diary of individual progress, successes and difficulties.
    b. In a unit, submit short reports on self at the culmination of difficult aspects of the work, and estimate the worth of the product and the benefits gained from the activity.
2. Conferences with the teacher can provide help, encouragement and direction.[21]

Trump was quite specific in his advocacy of student self-evaluation. He said, " . . . teachers must give pupils much more responsibility than is typically done today to assess their own progress in learning."[22] He stated that every classroom or independent study area needed a series of exercises prepared by teachers that would enable a pupil to

---

[20] Ibid., p. 84.

[21] Leonard H. Clark and Irving S. Starr, Secondary School Teaching Methods (New York: The Macmillan Company, 1967), p. 356.

[22] J. Lloyd Trump and Delmas F. Miller, Secondary School Curriculum Improvement (Boston: Allyn and Bacon, Inc., 1968), p. 353.

check his own progress. He used study in the English classroom to illustrate his method.

> For example in the field of English, exercises on capitalization, punctuation, number and tense, sentence structure, paragraph structure, and the like are provided. Pupils may work on these self-appraisal devices individually or with a partner. In any event, each pupil is encouraged to maintain his own personal records of what he has done.[23]

His advice to teachers included developing forms for student self-evaluations.

> One column can list the goals of the subject, defined in terms of what pupils are expected to do. A second column filled out by the learners can indicate approximately how well each student believes he will be able to accomplish the goals. A third column can indicate present progress as defined by each pupil with respect to the various goals. To report to parents and to the students, the teacher can then check, in the fourth column, his agreement or disagreement with the self-appraisal made by the student. A fifth column can be used for the pupil's final indication of agreement or disagreement.[24]

Trump believed that such a plan would be a tremendous move ahead in developing individual responsibility and understanding of the teaching-learning process.

To improve student self-evaluation techniques, Morsey prepared the following essay evaluation form to be used by an entire class to evaluate student papers.

---

[23] Ibid., p. 352.

[24] Ibid., p. 353.

ESSAY-EVALUATION FORM[25]

1 - Very poor
2 - Poor
3 - Below average
4 - Little below average
5 - Average
6 - Little above average
7 - Above average
8 - Good
9 - Very good
_ - Score

I. GRAMMAR AND USAGE (Subject-verb and pronoun-antecedent agreement, tense, case, distinction between adjectives and adverbs, etc.)

VP	P	BA	LBA	A	LAA	AA	G	VG
1	2	3	4	5	6	7	8	9

Score ____

II. Punctuation (Comma splice or comma fault, etc.)

VP	P	BA	LBA	A	LAA	AA	G	VG
1	2	3	4	5	6	7	8	9

Score ____

III. Sentence clearness and effectiveness (Sentence fragment, excessive coordination, faulty reference of pronouns, dangling modifiers, faulty parallelism, wordiness, etc.)

VP	P	BA	LBA	A	LAA	AA	G	VG
1	2	3	4	5	6	7	8	9

Score ____

IV. VOCABULARY (Right word, breadth, variety, adequacy, etc.)

VP	P	BA	LBA	A	LAA	AA	G	VG
1	2	3	4	5	6	7	8	9

Score ____

---

[25] Royal J. Morsey, Improving English Instruction (Boston: Allyn and Bacon, Inc., 1969), p. 284.

V. SPELLING

VP	P	BA	LBA	A	LAA	AA	G	VG
1	2	3	4	5	6	7	8	9

Score ____

VI. ORGANIZATION (Unity, coherence and emphasis in whole theme, etc.)

VP	P	BA	LBA	A	LAA	AA	G	VG
1	2	3	4	5	6	7	8	9

Score ____

VII. PARAGRAPHING (Unity, coherence, and emphasis within paragraph, etc.)

VP	P	BA	LBA	A	LAA	AA	G	VG
1	2	3	4	5	6	7	8	9

Score ____

VIII. CONTENT (Stimulating and fresh content, assumptions reasonable and supported with pertinent evidence, more than emotional reactions, etc.)

VP	P	BA	LBA	A	LAA	AA	G	VG
1	2	3	4	5	6	7	8	9

Score ____

TOTAL SCORE

Although Morsey stipulated that the essay-evaluation form was used in groups of four or five students, the form may be used by a single student evaluating himself. Morsey liked to compare his own evaluation of a student with that done by the group.

While Morsey submitted his essay-evaluation form to students, La Bernne was dealing with subjective self-reports and student participation in his assessments. La Bernne agreed that forms were helpful, but emphasized that questionaires, rating, scales, checklists and inventories were the

most frequently used in determining self-concept through introspective self-reflection and subjective self-reports. And he believed that the student should participate in self-evaluation.

> Evaluation should be a process of helping each student examine and expand his own skills. Here, the student participates in his own assessment based on his progress toward previously agreed-on goals. In this process the student can determine where he is and help chart the direction he may take.[26]

In looking at the questionnaire, rating scale, adjective checklist, or inventory the student and teacher could view how well the student completed his task. His goals and objectives could be contained in the form and self-evaluated as the student moved through a previously planned unit of work.

Morse and Wingo concluded that "The tendency toward self-evaluation should be encouraged and guided."[27] They believed learners could see where they needed to improve if they sought constructive self-evaluation. Record keeping was in order, according to these educators.

> One effective way in which the individual may evaluate himself is to keep a record of his own progress and samples of his work over a period of a semester or a year. In this way he has at hand concrete evidence of his

---

[26]Wallace D. La Bernne and Bert I. Greene, Educational Implications of Self-Concept Theory (Pacific Palisades, Calif.: Goodyear Publishing Company, Inc., 1969), p. 89.

[27]William C. Morse and G. Max Wingo, Psychology and Teaching (Glenview, Ill.: Scott, Foresman & Company, 1969), p. 491.

own development.[28]

Brian also supported those who advocated student self-evaluation on the part of the student. He concluded that evaluation by others predisposed an individual to resist; self-evaluation predisposes him to act. He stated, "It is essential that we do more to further self-evaluation, for this is the self-starter that energizes most programs of continued self-improvement."[29] At the outset in his remarks on evaluation, Brian showed discontent with the present grading process. He saw "nothing meaningful about a pupil expressed through our present marking system."[30] He believed that the system was no longer relevant to the needs and educational programs of our society. He noted that grades had no motivational effect and revealed nothing about what the student was in the process of becoming.

Combs was even stronger in his demand that schools train students to become self-directed learners. He emphasized, "Schools which do not produce self-directed citizens have failed everyone-the student, the profession, and the society they are designed to serve."[31] He stressed that the world we

---

[28]Ibid.

[29]George B. Brian, "Student Evaluation," The Education Digest, XXXIII, 4, (Dec., 1967), p. 53.

[30]Ibid., p. 52.

[31]Arthur W. Combs, "Fostering Self-Direction," Educational Leadership, Vol. 23, 5, (Feb., 1966), p. 373.

live in demands self-starting, self-directing students capable of independent action, and he made the point that responsibility and self-direction are learned.

Feedback and Self-Evaluations

In the literature several names emerged as strong advocates of student self-evaluations: Wilhelms, Grambs, Inlow, and Sawin. Sawin emphasized the importance of student self-evaluations by stating "self-evaluating is of crucial importance" and points to seeing "an increasing trend toward greater emphasis on the importance of both feedback and student self-evaluations."[32] He explained that it was difficult for one to improve on what he was doing unless he knew what would happen as a result of his efforts. Without this knowledge the student was likely to practice his errors and thereby reinforce them. Sawin stressed that students learned better if they were given information on the correctness of their efforts during the learning process.

He saw "the problem is not one of getting students to evaluate themselves; rather, it is getting them to do it accurately and in such a way as to contribute most effectively to their learning process."[33] He was convinced of the importance of feedback from the teacher and believed the teacher should explain evaluation results and materials most relevant

---

[32]Sawin, op. cit., p. 189.

[33]Ibid., p. 194.

to the student's current learning tasks.

"Feedback" was defined by Sawin as "the flow of information to the student on the correctness, adequacy, or appropriateness of his performance or reactions."[34] Sawin identified two parts to the process of providing feedback: evaluation and communication. In order to have information to communicate or to "feedback" to the student on his performance, it was first necessary to evaluate the performance to determine if it met expectations. He emphasized that evaluation was a part of the feedback procedure. Of great importance was the student's attitude toward the course of study and his own desire to improve.

> An important determiner of a student's actions in response to environmental stimuli is the student's own evaluation of his previous actions and their consequences - not the evaluation of the teacher or any other outsider but his own judgment on the extent to which his previous efforts represent progress toward his goals. Since the student's self-evaluation has important effects on his actions and since the actions determine what he learns, self-evaluation can be seen to have important effects on what he learns.[35]

Along these same lines, Sawin reasoned it seemed that, even though student self-evaluation might often be seriously inaccurate, it was an indispensable component in the students' learning processes. If a student's self-evaluations were inaccurate, it meant simply that the teacher had an important

---

[34] Ibid., p. 5.

[35] Ibid., p. 194.

task of helping the student learn better skills in self-evaluation so as to facilitate his learning, Sawin suggested.

Sawin believed the key to accuracy in student self-evaluations lay with the teacher's skill in developing student ability. A first requirement was to set forth the self-evaluation skills the student needed to develop. Self-evaluation in itself should be regarded as an important educational objective. Sawin suggested illustrations of kinds of objectives to improve skills in self-evaluation:

1. The student comprehends the fallacy of generalizing on the basis of a single instance. (An example is that success in solving one problem in long division does not mean that one can solve all problems in long division.)

2. The student is able to recognize when evidence is or is not relevant to the evaluation that he wishes to make.

3. The student understands that his own beliefs about what he can and cannot do may be inaccurate and that it often is necessary to get evidence obtained independently by outside observers.

4. The student comprehends the necessity of obtaining evidence on varied aspects of the ability or characteristic he wishes to evaluate. (Some students need to be reminded occasionally that they should not feel that they dislike all mathematics merely because they dislike what they studied in a particular course.)

5. The student comprehends the importance of focus and is able to achieve it in evaluating himself. (Students should select a few specific things about themselves to evaluate rather than attempt to evaluate many things at once.)

6. The student is able to make simple content analyses of his own work. (The student may make a content analysis of several of his arithmetic papers (and of the teacher's marks on them) to determine the types of multiplication problems on which he most commonly makes errors.)[36]

Sawin stipulated that students could probably evaluate themselves in some of these areas but not in the depth that teachers could. Teacher conferences and counseling would be a benefit here, he concluded.

Sawin was the first in the literature to suggest a point at which to begin self-evaluating. He suggested criteria for determining when it was wise to attempt self-evaluations.

1. A suitable goal, standard or criterion is available and understandable to the student.

2. It is possible for the student to obtain evidence which he understands and which is relevant.

3. The student genuinely wants to know the results of the evaluation. Self-evaluation cannot be forced. The student must accept the importance of the purpose to be served: the goals, criteria, or standards involved, and the validity of the evidence that will be used.

4. The results of the evaluation will not be harmful to the student. (There may be probability of making errors in interpreting the results of a test or task.)[37]

---

[36] Ibid., p. 196.

[37] Ibid., p. 197.

Sawin believed that certain conditions for learning had to be present to make for success in the learning process. He set the following objectives to be met to attain self-evaluation.

1. Set a good example in the practice of evaluation and self-evaluation. (The teacher should also practice self-evaluation and occasionally explain to students how he has improved his ability as a teacher by self-evaluation.)

2. Try to maintain a classroom atmosphere that encourages self-evaluation. Teachers can help create such an atmosphere by being more ready to say, "What do you think?" or "How well do you think you are progressing?" and not so ready to say, "The trouble with you is . . ." If the student is afraid to admit weaknesses and learning difficulties, this is likely to impair the effectiveness of the teacher in helping the student learn self-evaluation.

3. Conduct classroom activities in such a way that student efforts toward self-evaluation result in satisfying experiences.

4. Encourage self-evaluation, but do not try to force it, It has been previously explained that self-evaluation is possible only when the student wishes to know the results. From the principle of readiness, we can infer that students should be encouraged to make only those self-evaluations that they feel ready to make.

5. Give the student opportunities to practice the evaluation skills you want him to develop. Many teachers are inclined to try to do all the evaluating for the student. It should be self-evident, however, that a student will not learn a skill until he has an opportunity to practice it. This suggestion follows from that important principle of learning, the need for repetition.

6. Work at joint evaluation with students at least until they become able to comprehend what is involved in self-evaluation and are able to do it without help. Gradually turn more and more of the responsibilities for evaluation over to the students.

7. Have as many individual conferences for joint evaluation and planning as time permits, especially with students who are not doing well at self-evaluation.

8. At least until the student becomes rather skillful, discourage him from trying to evaluate too many things at a time.

9. Have students maintain written records of certain goals and their progress toward them.

10. Have class discussions for clarifying important goals and criteria for self-evaluation.

11. Have programmed textbooks, teaching machines, exercises with answers, mechanical devices (such as slide rule and an abacus for checking answers) available for supplementary use. These learning aids will provide enriched and very rapid feedback, which should facilitate self-evaluation.[38]

Sawin cautioned the teacher who was training the student in self-evaluations. He believed it was possible for the teacher to overwhelm students with too much feedback; he also thought some information could be harmful to the student's progress. He further warned teacher to be aware of inaccuracies likely to be characteristic of student self-evaluations. The student was still in the process of learning very complex

---

[38]Ibid., p. 199.

evaluation skills. Often social pressures and rewards in our society tend to lead to distortion of self-evaluations. Generally, conferences with the student produce corrections on his part and a more realistic self-evaluation.

Practicing Student Self-Appraisal

Long a proponent of student self-evaluations, Inlow believed teachers have engaged in student self-evaluations longer than they suspected. He observed that many teachers successfully injected pupil self-appraisal into the curriculum by asking pupils to evaluate their response to a unit being studied, how their behavior was modified as a result, and what they felt they needed to work on more specifically.

> Actually, every time a student consciously realizes an inadequacy and then works toward greater competency, he is engaging in self-evaluation. In the process of helping students toward this outcome, the teacher must provide helpful encouragement, accenting the positive and downgrading the negative, especially when the latter might be overly damaging to personality.[39]

Inlow had students evaluate themselves in several areas: knowledge of subject matter, ability to establish and maintain rapport with pupils, skill in discussion leadership, and tolerance under stress. He used a five-point rating continuum.

---

[39]Gail M. Inlow, Maturity in High School Teaching (Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1970), p. 318.

1 - very effective (or adequate)
2 - rather effective
3 - moderately effective
4 - rather ineffective
5 - very ineffective[40]

He assigned a value of one to five to a checklist of items, averaged the values, and then called the result an "outcome score." He found that "with greater academic competence comes greater accuracy in self-evaluation."[41] He did say, however, that the ultimate goal of an authority-centered evaluation in a school is pupil self-evaluation, and that self-evaluation in the less competent student is not lost. He stressed that "the ultimate goal of all evaluation is self-evaluation, which is a complex, lifelong process."[42]

The Student Self-Report Form

Grambs, mentioned in Chapter I, advocated the practice of developing self-evaluation tools. She emphasized that the teacher can give students a "technique for looking at their own achievement through a form developed jointly by the teacher and the class or by a class committee for the various phases or units of work."[43] When the student evaluates

---

[40]Ibid., p. 307.

[41]Ibid., p. 317.

[42]Ibid., p. 318.

[43]Grambs, op. cit., p. 337.

himself from the jointly developed form, student and teacher can appraise the results and discuss any differences that may arise. A variety of evaluative techniques may be used by the teacher: allow students to share in evaluating one another or use an evaluation committee with a constantly rotating membership, Grambs advised. Students then act in an advisory capacity in judging the quality of their own work. Grambs cited the following example.

> For example, it has been found very rewarding for small groups of students - three or four - to read and evaluate one another's papers. One teacher in a ninth grade orientation course had the students read their papers in small groups and then select the best one of the three or four read. After several periods of this kind of evaluation, students who had never written an acceptable paper or report were more conscientious in getting their papers done.[44]

The larger the role of the student in an evaluation program, the more effective total learning will be, Grambs suggested. Students developed in their ability to appraise and discriminate and then set their own standards and goals for achievement.

Grambs emphasized that her suggestions for overcoming some of the limitations of the grading process assume several operating principles.

> First, that grading is no longer the sole responsibility of the teacher, but becomes a joint endeavor in which the student takes an important part;

---

[44] Ibid., p. 338.

> Second, that each student can learn, can use his best talents, and can assess his own ability and effort;
>
> Third, that the actual achievement of the students is considered far more significant than the grade that is stamped upon it;
>
> Fourth, that continued teacher-student contact - many individual conferences, time spent on joint appraisal of progress, and cooperative planning of programs - is accepted as necessary.[45]

Student Involvement in Evaluation

Each year the Oklahoma State Department of Education publishes The Guide for Instruction in Communication. The 1970 issue asked teachers to "encourage rather than discourage" students when evaluating them. The Guide defined evaluation as "a continuous, cooperative, and cumulative process which includes pupil-teacher evaluation, pupil self-evaluation, and teacher evaluation."[46] Student self-evaluation activities include evaluating listening skills, speech competencies, reading competencies, and writing skills. The Guide included a list of questions for both teacher and student to answer in determining progress in learning. For example, the student may evaluate his own speech competencies with the following list of questions:

---

[45] Ibid., p. 340.

[46] A Guide for Instruction in Communication (Oklahoma State Department of Education, 1970), p. 45.

Do I feel at ease in conversations and discussions?

Do I look at everyone when I speak?

Do I disagree politely?

Do I try to bring others into the conversation?

Do I keep to the topic?

Do I choose and select words with care in order to convey exact meaning?

Do I enunciate distinctly?[47]

The student evaluated his own reading competencies by asking himself the following questions about selections:

Do I understand what I read?

Can I talk about what I read?

Do I feel that I am in the story when I read a short story or a novel?

Do I become upset, or emotionally involved when I read?

Do I know how a character feels?

The student had a chance to evaluate his own writing skills by answering the following questions:

Am I chiefly concerned with ideas in the first draft.

Can I express myself with clarity? - say what I want to say?

Do I use vivid and descriptive words to make my ideas come alive?

---

[47] Ibid., p. 46.

> Do I safeguard my meaning with good mechanics?
>
> Do I check the dictionary for spelling of words that are unfamiliar to me?
>
> Do I invite people to react to my ideas and to criticize my writing?[48]

The Guide provided questions to be answered by the teacher as she attempted to evaluate the student, but student self-evaluations were to be reviewed by the teacher, and conferences with the student were to be provided when differences in evaluations existed. The Oklahoma State Department placed the responsibility for evaluation with the teacher but allowed for the development of responsibility in the student. The Department stated quite clearly that in education "the goal is for self-evaluation."[49]

Peck concurred in placing responsibility for evaluation with the teacher but believed the student must be taught to be responsible.

> Young people want leadership, direction, some preparation for the unknown to come. We must ask students to evaluate both themselves and the system, but we must play our parts as the authorities. We must make very clear that we are responsible for their learning while they are learning to be responsible.[50]

Maughan, in a recent research study, placed the student

---

[48] Ibid., p. 48.

[49] Ibid., p. 45.

[50] Richard Peck, "Can Students Evaluate Their Education?" The Education Digest, Vol. XXXVI, 9 (May, 1971), p. 31.

in the forefront of the learning situation. He believed the student was capable of appraising his own performance.

> The pupil's own appraisal of the written performance affords his every advantage which makes for effective learning; in return the teacher avoids the pitfalls she would encounter were she to do the job alone. His own participation in the action of the appraisal takes the blindfolds off the learner and unshackles his fetters so that he truly learns.[51]

That learning and self-evaluation are interlinked seems to be the belief of many educators. Recently at least, that was what they seemed to be saying. Maughan stressed that the least that could be expected from student self-evaluation was greater proficiency in learning.

> If the premise is accepted as true - that self-evaluation is the professional approach which assures greater proficiency in learning from written endeavor, an evident conclusion can be reached - the learner is the one who must be in the action all the way.[52]

Pickup found "the procedure of asking a pupil for a written assessment of his own work was in itself a motivating influence;"[53] and advocated further research in this area. He stated that this idea was certainly worth investigating,

---

[51]Reese P. Maughan, "Learning Through Self-Involvement in the Evaluation of Written Effort," Education, Vol. 92, 4, (April-May, 1972), p. 64.

[52]Ibid., p. 65.

[53]A. J. Pickup and W. S. Anthony, "Teachers Marks and Pupils' Expectations: The Short-Term Effects of Discrepancies upon Classroom Performance in Secondary Schools," Social Psychology of Teaching (Baltimore, Maryland: Penguin Books, 1972), p. 342.

and the student had the opportunity to reflect about his own learning activity.

Eiszler conducted an investigation in student self-evaluations in an inner-city public school in a midwestern metropolitan area. His study involved 200 ninth grade black students, and found that accuracy in student self-evaluations were related to academic achievement and attitudes toward school.

> In the current study, task-specific self-evaluations are shown to be related to measures of academic achievement (grade point average), and a generally accepted measure of achievement motivation (estimates of performance). Based on the evidence of the current study, it may be tentatively concluded that task-specific self-evaluations are also related to a gross measure of attitude toward school (number of days absent).[54]

Eiszler further elaborated in his study that these results were interpreted as testimony to the validity of reported self-evaluations during task performance as a measure of achievement motivation. In other words, he found that the student did well if they liked school, attended, and showed some capabilities in their schoolwork.

Class Evaluations of Compositions

Hipple found class grading of various classmates' papers

---

[54]Charles F. Eiszler and Betty M. Morrison, "Task-Specific Self-Evaluation as a Measure of Achievement Motivation," The Journal of Experimental Education, Vol. 40, 3 (Spring, 1972), p. 31.

in composition was rewarding and a valid learning situation in an English class. He suggested that if knowing a student's name when his paper was being evaluated posed a problem, the paper should be coded. He believed that total class grading of the same paper can alert the students to the difficulties inherent in so subjective a process as composition evaluation. Sometimes the students assumed defensive positions as they explained how they arrived at the grades they gave.

> More fruitful, however, is their learning about how to evaluate, how to choose certain errors to stress even while they ignore others, and, most important, how to suggest that the written product be improved.[55]

Hipple believed that this type of training by doing aided the student in learning to do his own self-evaluation. The experience in evaluating others and the comparison with his own composition was valuable training in self-evaluation. Hipple did not discount the practice of letting the student name the grade he thought he deserved, for the practice was not without merit, he concluded.

> The student, after all, is probably the best judge of his effort and accomplishment. His being his own grader may enable him to learn something of self-evaluation, a vital component of one's becoming a mature individual. With some students, too, the knowledge that they will be grading themselves serves as a motivating force, a push toward the best work of which they are capable if only to forestall any ambivalence about grades they award themselves.[56]

---

[55]Theodore W. Hipple, Teaching English in Secondary Schools (New York: The Macmillan Company, 1973), p. 164.

The Conservative Approach to Student Self-Evaluations

In the literature related to student self-evaluations, there were as many techniques for student self-evaluations as there were proponents of self-evaluations. The beginning teacher or the skeptical teacher may be more comfortable with the techniques of Alexander and Halverson, early advocates of student self-evaluations. These educators took the moderate road and emphasized goal-setting on the part of the student wishing to learn to evaluate himself. They set the following steps to be followed in the student's own evaluation.

1. to define the goal being sought
2. to collect evidence on the realization of the goal
3. to make a judgment as to whether the achievement is satisfactory
4. in light of the value determined, to plan next steps[57]

Alexander and Halverson pointed out that these steps took on more specific meanings when the concern was evaluation of progress of students. They suggested that each student established for himself some behavior to be achieved. Since emphasis was placed on progress or improvement, there was a need for taking stock of present behavior which could be compared later with behavior of that time. They

---

[56] Ibid., p. 291.

[57] Alexander and Halverson, op. cit., p. 399.

emphasized that the significant step then was that of defining for each pupil, if possible, by the pupil himself, the kind of behavior desired.

> Once the desired behavior or goal is defined, evidence of progress may be collected and a record kept of that progress. In this way a student can note the extent of his progress in each task completed.[58]

Alexander and Halverson believed that possibly the most effective evaluation was done wholly informally as the pupil saw that he had behaved or not behaved in a specific instance in the way desired. Alexander and Halverson concluded that traditional testing by the teacher was a necessary part of the total evaluation scheme. Other areas had to be evaluated, however. Some activities were observed; some were performed in group situations and some were oral discussions. Whatever the basis and time for systematic judgment making, the process was uniformly that of looking at the evidence of progress in relation to the original prediction as to the type of behavior desired and deciding whether the progress was satisfactory, these educators maintained.

Most desired behavior required periodic appraisal and judgment and planning for additional steps, Alexander and Halverson stated. They saw the student and teacher achieve this step in group situations and personal conferences.

Regardless of the procedure of collecting

---

[58] Ibid., p. 400.

> evaluative evidence, all effective evaluation becomes self-evaluation. That is, pupil progress is ultimately motivated by the pupil's own planning, and this planning has to be based on some type of evaluative evidence. This evidence may come through the pupil's own processes, or he may accept - and acceptance is itself an act of self-evaluation - the evidence from his teacher, the learning group, or others.[59]

Ideally, the student did not rely on the teacher alone or on the group alone, but he evaluated himself, as he should, by honestly and accurately looking at the evidence, his own progress, and his own learning process.

In conclusion, student self-evaluations were recognized by many educators as a natural outgrowth of a student's own learning development. Teachers' grades were not always the answer to the evaluative process, although they were needed. Teachers could teach students how to evaluate themselves honestly and accurately. Student self-evaluations were a part of any course of study, these educators believed, and a necessary part of developing student responsibility. The techniques of student self-evaluations to be used varied, and most were quite adaptable to any course of study. The next step lies with the individual teacher to incorporate the practice of student self-evaluations into the course of study. First attempts may not be perfect, but the student can learn to evaluate himself and appreciate the value in

---

[59] Ibid., p. 400.

doing so. Educational theories regarding student self-evaluations did exist in the literature, but actual studies implementing these theories were not plentiful.

CHAPTER III

METHODS AND PROCEDURES

In the present study two groups of students enrolled in four (4) high school English classes were compared on their ability to conduct self-evaluations of their weekly progress and overall class performance. Two of the classes were trained to conduct self-evaluations of their progress in five course units (Unit Topics). Using the Student's Self-Evaluation Rating Sheet shown in Appendix B, these two classes (N_1=32; N_2=20; Total N=52) made weekly progress ratings on a five-point continuum. The second group, classes three and four, were given no training in self-evaluation techniques but were asked to approximate the final grade they would receive in the class. The two groups which received no self-evaluation training (N_3=25; N_4=25; Total N=50) approximated the final grade they would receive in the course after the five units of study had been completed. The course grades estimated by the two groups of participants were compared with the grades assigned by the course instructor to test the hypotheses stated in the study.

The methods and procedures actually used in the study were classified into the following three phases: (1) Pre-Experimental Procedures, (2) Experimental Procedures and (3) Data-Analysis Procedures. Each of these phases is discussed in detail in the following sections.

PRE-EXPERIMENTAL PROCEDURES

The pre-experimental procedures consisted of all those tasks which the researcher had to complete before the actual collection of the data began. The most important of these tasks were described in the following sections.

Choice of Research Design

The first pre-experimental procedure was to choose the proper research design for the conduct of the study. The words "research design" are intended to mean the plan, structure, and strategy of investigation conceived to obtain answers to research questions and to control external sources of variation. The Plan is the overall scheme or program of the evaluation problem; the Structure is the more specific structure or paradigm of the actual manipulation of the independent variables being controlled; and the Strategy as used here is even more specific than the structure--it is the actual methods to be used in the gathering and analysis of the data.

A research design serves two basic purposes: (1) it provides answers to research questions posed by the investigator; and (2) it controls external sources (independent variables) of variation. Without the proper research design the results of any investigation would be meaningless. F. N. Kerlinger made the following statement in regard to the research and evaluation designs:

> . . . How does design accomplish this? Research designs set up the framework for 'adequate' tests of the relations among variables. The design tells

> us, in a sense, what observations (measurements) to make, how to make them, and how to analyze the quantitative representations (data) of the observations. Strictly speaking, design does not 'tell' us precisely what to do, but rather suggests the directions of observation-making and analysis, how many observations should be made, and which variables (independent variables) are active variables and which are assigned. We can then act to manipulate (control) the active variables and to dichotomize or trichotomize or otherwise categorize the assigned variables. A design tells us what type of statistical analysis to use. Finally, an adequate (proper for the particular situation) design outlines possible conclusions to be drawn from the statistical analysis (Parentheses material added).[1]

The research design chosen for the present experiment was a multiple-sample true experimental design preceded by the random sampling of participants from four (4) finite populations. A paradigm of this research design is presented in Figure 1.

Development of a Self-Evaluation Rating Sheet

The questionnaire rating instrument used in this study was developed in close cooperation with the teacher, students from the four classes, and the Chairman of the Department of English. The purpose of the questionnaire was to assist the student in making a fair and objective appraisal of his progress during a particular Unit Topic.

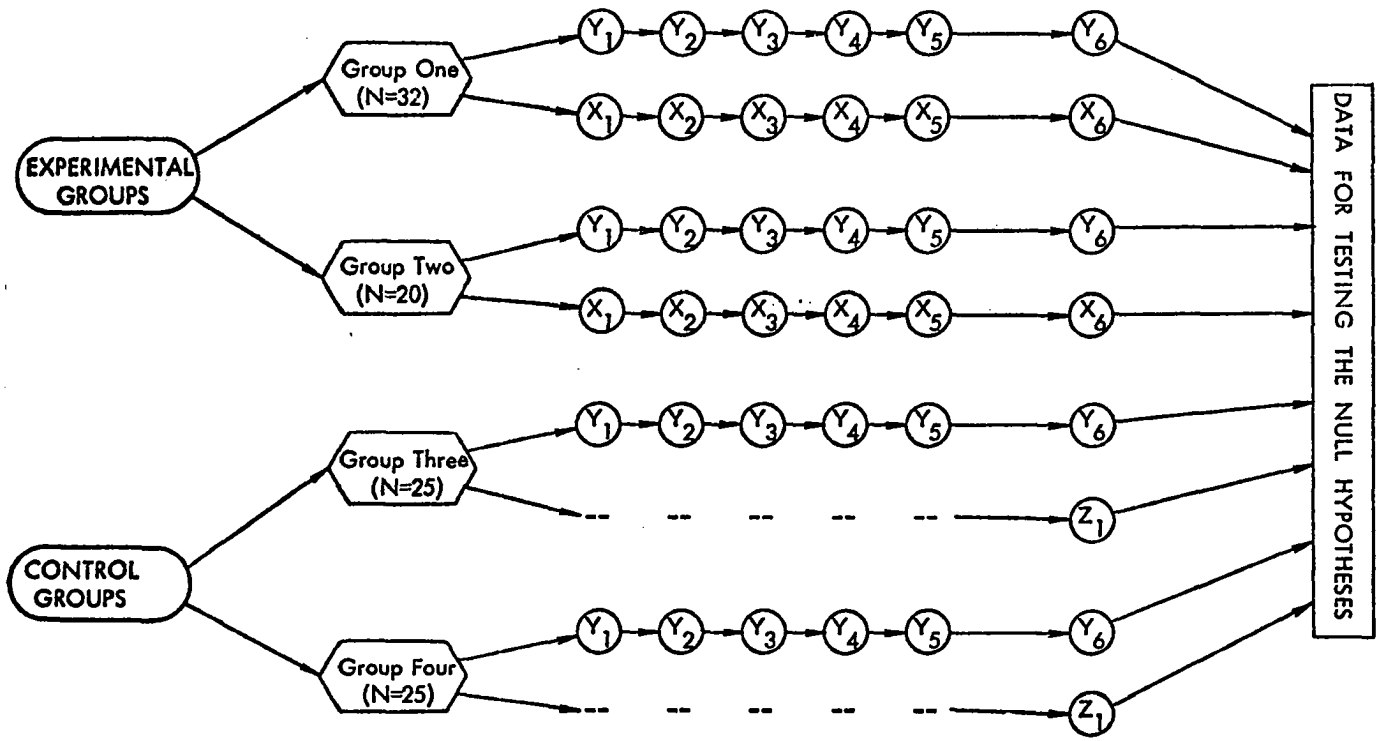
The first step in accumulating items for a rating sheet

---

[1]Fred N. Kerlinger, Foundations of Behavioral Research, (New York: Holt, Rinehart and Winston, 1964), pp. 196-97.

Figure 1

RESEARCH DESIGN USED IN THE EXPERIMENT



Explanation of Symbols:

Experimental Groups: Those students who were taught self-evaluation techniques

Control Groups: Those students who were not taught self-evaluation techniques

Y_{1-5} : Unit Topic grades assigned by the course instructor

Y_6 : Final course grade assigned by the course instructor

X_{1-5} : Unit Topic grades estimated by the experimental group students

X_6 : Final course grade estimated by the experimental group students

Z_1 : Final course grade estimated by the control group students

for student self-evaluations was to set objectives for the five-unit study project which covered five weeks of reading, answering questions, discussing the literature, essay writing, and testing. The teacher submitted general objectives (traditional in a literature class) to the experimental groups for their consideration and also to serve as a guide. The objectives were discussed in each class. Students added objectives in their discussion groups; these were discussed; some were dropped by common consent, and others were added. A final set of objectives for the study was adopted. (see Appendix A).

A similar procedure was followed in developing the Student Self-Evaluation Rating Sheet. Students in the experimental groups were given a fifteen-item (15) rating sheet with traditional literature study items included. Students discussed the items (No new items were added at this time), and rated themselves on a literary analysis of several short stories by Hawthorne. The ratings continued over a period of five weeks with specified assignment tests, open discussions, and private counseling at each weekly rating session. The five weeks of pretesting indicated that the instrument could be improved by adding some items.

After five weeks of students evaluating themselves on the fifteen-item rating sheet, the teacher submitted the ratings to the experimental groups for discussion, evaluation, and adjustment. The twenty-four item rating sheet evolved

from the meetings. It was compiled in its final form by the classes of English students and edited by the teacher. These twenty-four items comprised the final Student's Self-Evaluation Rating Sheet for a weekly rating of progress in the five-unit literature study.

Establishing the Reliability and Validity of the Data Collection Instrument

To establish the reliability of the questionnaire, it was necessary to pretest the instrument before adapting the final form. The following ideas are recommended by Good:

> Before the final form (of the questionnaire) is prepared and distributed to the respondents, tryout of pretesting of the questionnaire is essential, for the purpose of validation in terms of practical use. This tryout probably will lead to revisions of certain questions, deletion of useless questions, and addition of other items. Tabulation of the tryout responses in rough tables will indicate whether the answers can be tabulated satisfactorily and whether answers to the major questions are forthcoming. The manual of the United States Bureau of the Budget emphasizes that it is desirable to test the feasibility of the questionnaire survey in advance, with pretests designed and conducted to secure answers to such problems as the following:
>
> (a) Relative effectiveness and costs of alternative questionnaires, instructions, and operating procedures.
>
> (b) Acceptability and intelligibility of the questions from the respondent's point of view.
>
> (c) Possible misunderstandings of questions and procedure on the part of the interviewers.
>
> (d) Clarity and applicability of definitions and classifications.

(e) Completeness of questions for correct coding and interpretation.

(f) Defects in the forms, maps, lists, instructions, etc.

(g) Estimates of strata means and variances.

(h) Response rates.[2]

A final effort to establish the validity of the instrument was made through personal observations of the students as they completed the questionnaire.

Training the Students to Conduct Self-Evaluations of Their Progress

The next step in the pre-experimental procedures was to train the two experimental groups in self-evaluation procedures. This training was actually a continuing process and was not completed prior to the collection of the data. After the first Unit Topic was completed the course instructor assigned each student a Unit Topic grade. After returning these grades, the course instructor distributed the self-evaluation rating sheets to the two experimental groups (Group One and Group Two). Students in these classes were asked to complete these rating instruments based on their progress during the week of the first Unit Topic "Play Analysis". Student self-evaluations were performed after each of the five Unit Topics was completed. The completion

---

[2]Carter V. Good, Essentials of Educational Research, Methodology and Design, (New York: Appleton, Century, Crofts, 1966), p. 262.

of these rating instruments on a periodic basis constituted the majority of the self-evaluation training given to the experimental groups.

Each time the Student Self-Evaluating Rating Sheet was submitted to the experimental group, a progress discussion was conducted concerning any problem students had rating themselves. For example, some students hesitated to rate themselves on item #15 if they were rating study questions. They were instructed to consider study question answers under the same category as essay writing, since the structure rules for writing were generally the same. They were also expected to observe the rules of grammar, punctuation, and spelling in whatever they wrote.

Some students needed counseling to complete their individual rating sheets, but most improved their self-evaluating abilities as they continued the week-to-week ratings. Most advantageous to the experimental group was the weekly reminder of progress, and many benefited from the sessions which focused on individual improvement.

Selection of the Participants

The next step of the pre-experimental procedures was the selection of the study participants. It was not possible to randomly assign students to the experimental and control classes, and the researcher was forced to randomly choose two classes as experimental groups (those who were taught self-evaluation techniques) and two classes as control groups

(those who were not taught self-evaluation techniques). The sampling paradigm is shown in Figure-2.

EXPERIMENTAL PROCEDURES

The experimental procedures consisted of all those tasks which were completed in the actual collection of the data from the four groups of participants. These procedures are shown in Figure 1.

There were two basic measures reported by students in the two experimental groups. (1) a total average progress rating (grade) based on the five Unit Topic ratings and (2) the class grade assigned by the course instructor. The five units of literature studied in the experiment were as follows:

1. Unit One: Play Analysis; identified the study unit of a three-act play. After reading and discussing the play, students wrote an essay concerning the themes of the play. The essay was evaluated by the teacher on a grading basis of A, B, C, D, or F and returned to the students. They, then, corrected their errors and rated their progress on the Student Self-Evaluation Rating Sheet. Most were fairly proficient at rating themselves by this time, since they had the earlier experience of rating themselves on the fifteen-item rating sheet. However, some required individual counseling and further explanation of the twenty-four items on the new rating sheet.

2. Unit Two: American Humor; identified the study unit of an assortment of American humorists' writings. Students

Figure 2

SAMPLING PARADIGM OF THE EXPERIMENT

EXPERIMENTAL GROUPS	SECTION I	SECTION II	TOTAL SECTIONS
	(N = 32)	(N = 20)	(N = 52)
CONTROL GROUPS	SECTION I	SECTION II	TOTAL SECTIONS
	(N = 25)	(N = 25)	(N = 50)
TOTALS . . .	SECTION I	SECTION II	GRAND TOTALS
	(N = 57)	(N = 45)	(N = 102)

read the assigned materials, completed sets of pre-written questions, and submitted their answers for grading. Responses were graded by the teacher and returned to students for corrections. A discussion followed, and students in the experimental groups were asked to rate their performance on the unit. Again, some students sought further individual counseling, some discussed their progress, and others asked for individual help in improving future performances.

3. Unit Three: Novel Analysis; identified the literature study of a novel. Students were allowed to select any novel they wished as long as it was written by a noted American author. They received instructions in analyzing plots, characters, themes, moods, and settings, and were asked to analyze a novel. The students' analyses were evaluated by the teacher and returned for corrections and student self-evaluations.

4. Unit Four: The Twain-Thurber Study Guide; consisted of a set of questions over these two writers. After students had been briefed on the backgrounds of these writers, their objectives in writing, their general meanings in American life, and their impact on society, they were asked to complete sets of study guide questions. Student responses were evaluated by the teacher and returned for their corrections and self-evaluation.

5. Unit Five: Naturalism in Contemporary Fiction; included short stories by Hemingway, Bradbury, Steinbeck,

Faulkner, and Stegner. Students wrote essays discussing the characteristics of naturalism detected in the short stories read. Essays were submitted for grading; writings were evaluated and returned to the students for corrections. They then evaluated their performances on the rating sheet.

The average progress rating of the five Unit Topics was computed in the following steps:

1. The five Unit Topic ratings from each of the twenty-four questionnaire items were averaged.
2. Numerical averages for the twenty-four questionnaire items were again averaged to arrive at a total average rating for each student.
3. Total average ratings were considered to be the average progress ratings assigned by the students.

Total average ratings shown as $\overline{X}_6$ in Figure 1, were collected for the two experimental groups. Class grades assigned by the course instructor, shown as Y_6 in Figure 1, were based on each student's performance in the course.

Students in the two control groups also reported two evaluation measures during the course of the experiment. They were asked to approximate the grade they would receive in the five unit study, and they were assigned a grade by the course instructor. The control groups made their grade approximations after the five Unit Topics had been completed.

After all data had been collected from both groups, the preliminary analysis began. The tasks performed in these analyses are presented in the data analysis section.

DATA-ANALYSIS PROCEDURES

The data-analysis procedures consisted of those tasks performed after the data had been collected from the participants. These procedures included the coding of the data, selection of statistical procedures, and testing of the hypotheses.

The first step in the data-analysis procedures was the coding of the participants' responses. It was necessary to assign numerical values to the grades approximated by the two control groups and the grades assigned by the course instructor. The numerical values assigned to the various letter grades were as follows:

A - 4.00
B - 3.00
C - 2.00
D - 1.00
F - 0.00

No provision was made in the grade coding for a minus (-) or plus (+) grade. While this simplified the coding process, part of the grading information was lost. Assigning numerical values to letter grades and averaging the Unit Topic ratings resulted in two measures for each participant.

Choice of Statistical Procedures

The next step of the data analysis procedures was the selection of the proper statistical procedures to be used in testing the hypotheses. Several criteria were taken into consideration when choosing the testing statistics. These criteria were as follows:

(1) The nature of the hypothesis being tested;

(2) The measurement level of the data being used in the comparison;

(3) The number of groups being compared at any one time;

(4) The number of participants within each group being compared;

(5) The assumptions underlying the statistical test(s) chosen.

The statistical tests chosen for each of the five hypotheses were screened according to the five criteria listed. The statistical tests chosen in each case are shown in Figure 3 along with the null hypothesis being tested and the measures (data) needed to make the comparison.

Figure 3

INFERENTIAL STATISTICS CHOSEN TO TEST EACH NULL HYPOTHESIS

Hypothesis Number	Null Hypothesis Being Tested		Testing Statistic(s)	Data Involved In The Calculations
One	Ho_1	There is no statistically significant difference between the course grade estimates made by high school students who are taught self-evaluation techniques and the course grades assigned to them by their instructor.	t test for two dependent (correlated) measures	(1) Experimental students' grade estimates (2) Course grades assigned by the instructor
Two	Ho_2	There is no statistically significant difference between the course grade estimates made by high school students who are not taught self-evaluation techniques and the course grades assigned to them by their instructor.	t test for two dependent (correlated) measures	(1) Control groups' course grade estimates (2) Course grades assigned by the instructor
Three	Ho_3	There is no statistically significant difference between the estimated/actual course grade differences of high school students who are taught self-evaluation techniques AND the estimated/actual course grade differences of high school students who are not taught self-evaluation techniques.	t test for two independent (uncorrelated) measures	(1) Estimated/actual grade differences of experimental gp. (2) Estimated/actual grade differences of control group
Four	Ho_4	There is no statistically significant relationship between the course grade estimates made by high school students who are taught self-evaluation techniques and the course grades assigned to them by their instructor.	Pearson Product-Moment Correlation Coefficient "r"	(1) Experimental students' grade estimates (2) Course grades assigned by the instructor
Five	Ho_5	There is no statistically significant relationship between the course grade estimates made by high school students who are not taught self-evaluation techniques and the course grades assigned to them by their instructor.	Contingency Coefficient "C"	(1) Control groups' course grade estimates (2) Course grades assigned by the instructor
Six	Ho_6	There is no statistically significant difference between the correlation coefficient computed between the estimated and actual course grades of the two experimental groups (those who were taught self-evaluation techniques) AND the correlation coefficient computed between the estimated and actual course grades of the two control groups (those who were not taught self-evaluation techniques).	z test between two independent correlation coefficients	(1) Correlation coefficient computed in hypothesis four (2) Correlation coefficient computed in hypothesis five

CHAPTER IV

RESULTS OF STATISTICAL ANAYSIS

Data collected from 102 students (four classes) enrolled in high school English courses were used to test six null hypotheses. Two of the classes were designated as experimental groups, and two were designated as control groups. The two experimental groups (N_1=32; N_2=20; Total N=52) used a Student's Self-Evaluation Rating Sheet to make weekly ratings of their progress in five course units (Unit Topics). The course grades estimated by the experimental groups were compared to the actual course grades assigned by their instructor. The two control groups (N_1=25; N_2=25; Total N=50) also were asked to approximate the grade they would receive in the course. However, these two groups did not receive the training in self-evaluation techniques which was being given to the experimental groups. The researcher had hypothesized that the self-evaluation training would enable the two experimental groups to make significantly more accurate estimates of the course grades they would actually receive than the two control groups. The results of testing the six null hypotheses stated in Chapter I are presented in this Chapter of the dissertation. A summary of the study, the conclusions drawn from the results, and implications for further research are presented in Chapter V.

Descriptive Statistics of the Four Classes

The descriptive data computed for the four classes are presented in Table 1. This Table contains three measures for each student participant; (1) the course grade/score estimated by the student, (2) the course grade/score assigned by the instructor, and (3) the arithmetic difference between the estimated and acutal grade/scores. The means and standard deviations computed for each class and each group are presented in Table 1. The raw data are presented in the Appendices along with the appropriate descriptive statistics.

Results of Testing the Hypotheses

The results of testing the six null hypotheses are presented in the following sections of the dissertation. In each instance the exact null hypothesis is stated, followed by a Table containing the statistical data and a brief explanation of the results. Additional results are presented in the ancillary findings section.

Results of Testing the First Null Hypothesis Ho_1

The null proposition of the first null hypothesis was tested as follows:

Ho_1 There is no statistically significant difference between the course grade estimates made by high school students who are taught self-evaluation techniques and the actual course grades assigned to them by their instructor.

The first null hypothesis was tested by comparing the course grade estimates made by the students who had been

TABLE 1

DESCRIPTIVE STATISTICS OF GRADES ESTIMATED BY STUDENTS, GRADES ASSIGNED BY THE COURSE INSTRUCTOR, AND THE ARITHMETIC DIFFERENCES

	SECTION I			SECTION II			TOTAL GROUPS		
	Estimated	Actual	Difference	Estimated	Actual	Difference	Estimated	Actual	Difference
EXPERIMENTAL GROUPS	$\bar{X}$ = 2.154 sd = 1.047 N = 32	$\bar{X}$ = 2.000 sd = 1.136 N = 32	$\bar{X}$ = -0.154 sd = 0.515 N = 32	$\bar{X}$ = 2.816 sd = 0.558 N = 20	$\bar{X}$ = 3.000 sd = 0.725 N = 20	$\bar{X}$ = 0.185 sd = 0.536 N = 20	$\bar{X}$ = 2.408 sd = 0.942 N = 52	$\bar{X}$ = 2.385 sd = 1.105 N = 52	$\bar{X}$ = -0.024 sd = 0.544 N = 52
	Estimated	Actual	Difference	Estimated	Actual	Difference	Estimated	Actual	Difference
CONTROL GROUPS	$\bar{X}$ = 2.280 sd = 1.021 N = 25	$\bar{X}$ = 2.080 sd = 1.077 N = 25	$\bar{X}$ = -0.200 sd = 0.500 N = 25	$\bar{X}$ = 2.360 sd = 0.995 N = 25	$\bar{X}$ = 1.680 sd = 1.108 N = 25	$\bar{X}$ = -0.720 sd = 0.936 N = 25	$\bar{X}$ = 2.320 sd = 0.999 N = 50	$\bar{X}$ = 1.880 sd = 1.100 N = 50	$\bar{X}$ = -0.460 sd = 0.788 N = 50
	Estimated	Actual	Difference	Estimated	Actual	Difference	Estimated	Actual	Difference
TOTALS	$\bar{X}$ = 2.209 sd = 1.029 N = 57	$\bar{X}$ = 2.035 sd = 1.101 N = 57	$\bar{X}$ = -0.174 sd = 0.504 N = 57	$\bar{X}$ = 2.562 sd = 0.852 N = 45	$\bar{X}$ = 2.267 sd = 1.156 N = 45	$\bar{X}$ = -0.318 sd = 0.899 N = 45	$\bar{X}$ = 2.363 sd = 0.972 N = 102	$\bar{X}$ = 2.137 sd = 1.126 N = 102	$\bar{X}$ = -0.238 sd = 0.706 N = 102

Estimated = Grades/Scores estimated by the students
Actual = Grades actually assigned by the course Instructor
Difference = Arithmetic difference between estimated and assigned grades

$\bar{X}$ = Mean (Average)
sd = Standard Deviation
N = Number of grades/scores within the subgroup

taught self-evaluation techniques and the actual course grades assigned to them by their instructor. The means of these two sets of data were compared with a student's t test for correlated data. The results of the statistical calculations are presented in Table 2 along with the means and standard deviations used to make the comparison.

The results presented in Table 2 show that there was no significant difference between the course grades estimated by the students in the experimental groups and the course grades assigned by the instructor. These results tend to support the general hypothesis that teaching students self-evaluation techniques enables them to make accurate approximations of the grade they will receive in a particular course.

Results of Testing the Second Null Hypothesis Ho_2

The null proposition of the second null hypothesis was tested as follows:

Ho_2 There is no statistically significant difference between the course grade estimates made by high school students who are not taught self-evaluation techniques and the actual course grades assigned to them by their instructor.

The second null hypothesis was tested by comparing the course grade estimates made by the students who had not been taught self-evaluation techniques and the actual course grades assigned to them by their instructor. The means of these two sets of data were compared with a student's t test for correlated data. The results of the statistical calculations

TABLE 2

A COMPARISON OF COURSE GRADE ESTIMATES MADE BY THE EXPERIMENTAL GROUPS AND THE ACTUAL COURSE GRADES ASSIGNED BY THE INSTRUCTOR

Source of Grade/ Score Values	Mean	Standard Deviation	t-Value	Significance Level
Grades Estimated by Experimental Groups	2.4084	0.9423		
			0.3126	$p > .05$
Grades Assigned by Course Instructor	2.3846	1.1053		

are presented in Table 3 along with the means and standard deviations used to make the comparison.

The results presented in Table 3 show that there was a significant difference between the course grades estimated by the students in the control groups and the course grades assigned by the instructor. These results tend to support the general hypothesis that unless students are taught self-evaluation techniques in a particular course they are unable to make accurate approximations of the grades they will receive from the instructor of that course.

Results of Testing the Third Null Hypothesis Ho_3

The null proposition of the third null hypothesis was tested as follows:

Ho_3 There is no statistically significant difference between the estimated/actual course grade differences of high school students who are taught self-evaluation techniques and the estimated/actual course grade differences of high school students who are not taught self-evaluation techniques.

The third null hypothesis was tested by comparing the magnitude of the errors made by the experimental groups in estimating their course grades with the magnitude of the errors made by the control groups in estimating their course grades. The means of the two sets of difference scores were compared with a *t* test for two independent means. The results of the statistical calculations are presented in Table 4 along with the means and standard deviations used to make the comparison.

TABLE 3

A COMPARISON OF COURSE GRADE ESTIMATES MADE BY THE CONTROL GROUPS AND THE ACTUAL COURSE GRADES ASSIGNED BY THE INSTRUCTOR

Source of Grade/ Score Values	Mean	Standard Deviation	t-Value	Significance Level
Grades Estimated by Control Groups	2.3200	0.9987		
			2.092	$p < .05$
Grades Assigned by Course Instructor	1.8800	1.1000		

The results presented in Table 4 show that there was a significant difference between the grade estimation errors committed by the experimental groups (those who had been taught self-evaluation techniques) and the grade estimation errors committed by the control groups (those who had not been taught self-evaluation techniques). While both groups tended to underestimate the grades they would receive in the course, students from the experimental groups were significantly more accurate in their estimations than students from the control groups.

Results of Testing the Fourth Null Hypothesis Ho_4

The null proposition of the fourth null hypothesis was tested as follows:

Ho_4 There is no statistically significant relationships between the course grade estimates made by high school students who are taught self-evaluation techniques and the course grades assigned to them by their instructor.

The fourth null hypothesis was tested by computing a Pearson's Product-Moment correlation coefficient ("r") between the grades estimated by the experimental groups and the grades assigned these students by the course instructor. The means ($\overline{X}$) and standard deviations (sd) of these two sets of data are presented in Table 5 along with the resulting correlation coefficient and the significance level of the results.

The statistical results presented in Table 5 indicate that there was a significant relationship between the grades

TABLE 4

A COMPARISON OF THE ESTIMATED/ACTUAL COURSE GRADE DIFFERENCES COMPUTED FOR THE EXPERIMENTAL AND CONTROL GROUPS

Source of Estimated/ Actual Course Grade Differences	Mean	Standard Deviation	t-Value	Significance Level
Students in Experimental Groups (N=52)	-0.0238	0.5437		
			3.2426	$p < .001$
Students in Control Groups (N=50)	-0.4600	0.7879		

TABLE 5

THE RELATIONSHIP BETWEEN THE COURSE GRADE ESTIMATES MADE BY THE EXPERIMENTAL GROUPS AND THE ACTUAL COURSE GRADES ASSIGNED BY THE INSTRUCTOR

Source of Grade/ Score Values	Mean	Standard Deviation	Correlation Coefficient	Significance Level
Grades Estimated by Experimental Groups	2.4084	0.9423		
			$r = 0.8708$	$p < .001$
Grades Assigned by Course Instructor	2.3846	1.1053		

estimated by the students who were taught self-evaluation techniques and the grades assigned to them by the course instructor. These results tend to support the results derived from testing the first null hypothesis (Table 2). The magnitude of the correlation coefficient allowed the researcher to reject the fourth null hypothesis.

Results of Testing the Fifth Null Hypothesis Ho_5

The null proposition of the fifth null hypothesis was tested as follows:

Ho_5 There is no statistically significant relationships between the course grade estimates made by high school students who are not taught self-evaluation techniques and the course grades assigned to them by their instructor.

The fifth null hypothesis was tested by computing a Contingency Coefficient ("C") between the grades estimated by the two control groups and the grades assigned these students by the course instructor. The type of correlation coefficient used to test the fifth null hypothesis was different from that used to test the fourth null hypothesis because of the different measurement levels of the data involved. The data involved in the fourth hypothesis showed one variable as being continuous (grades estimated by the students) and one variable as being continuous but forced into a polychotomy (grades assigned by the course instructor). This required a Pearson Product-Moment Correlation Coefficient

("r").[1] However, the data involved in the fifth hypothesis showed both variables as being continuous but forced into dichotomies. The proper correlation technique in this instance was a Contingency Coefficient ("C").[2] The results of the statistical calculations are presented in Table 6 together with the means and standard deviations of both sets of data.

The statistical results presented in Table 6 indicate that there was no significant relationship between the grades estimated by the students who were not taught self-evaluation techniques (control group) and the grades assigned to them by the course instructor. Again, the results tended to support the results derived from testing the second null hypothesis presented in Table 3

Results of Testing the Sixth Null Hypothesis Ho_6

The null proposition of the sixth null hypothesis was tested as follows:

Ho_6 There is no statistically significant difference between the correlation coefficient computed between the estimated and actual course grades of the two experimental groups (Those who were taught self-evaluation techniques) AND the correlation coefficient computed between the estimated and actual course grades of the two control groups (Those who were not taught self-evaluation techniques).

---

[1] N. M. Downie and R. W. Heath, Basic Statistical Methods, (New York: Harper and Row, 1965), pp. 78-94.

[2] Ibid. pp. 210-212.

TABLE 6

THE RELATIONSHIP BETWEEN THE COURSE GRADE ESTIMATES MADE BY THE CONTROL GROUPS AND THE ACTUAL COURSE GRADES ASSIGNED BY THE INSTRUCTOR

Source of Grade/ Score Values	Mean	Standard Deviation	Correlation Coefficient	Significance Level
Grades Estimated by the Control Groups	2.3200	0.9987		
			C = 0.0401	p >.05
Grades Assigned by the Course Instructor	1.8800	1.1000		

The sixth null hypothesis was tested by comparing the two correlation coefficients computed in hypothesis four and five. This comparison was made by transforming the correlation coefficients to a z format via a Fisher's z Transformation, and computing a z test for two independent correlations.[3] The results of the statistical calculations are presented in Table 7 along with the correlation coefficients, the z transformations, and the significance level of the results.

The statistical results presented in Table 7 indicate that there was a significant difference between the correlations computed for the two groups of students. There was a significantly stronger relationship between the instructor's course grades and the grades estimated by the two groups of experimental students (those who had been taught self-evaluation techniques) than there was between the instructor's course grades and the grades estimated by the two groups of control students (those who had not been taught self-evaluation techniques). These results allowed the researcher to reject the sixth null hypothesis and conclude that the self-evaluation techniques taught to the two experimental groups proved beneficial to them when they were attempting to approximate their course grades. Conclusions drawn from the findings are located in the final chapter of this paper.

---

[3]James L. Bruning and B. L. Kintz, Computational Handbook of Statistics, (Dallas, Texas: Scott, Foresman and Co., 1968), pp. 193-194.

TABLE 7

A COMPARISON OF THE CORRELATION COEFFICIENTS COMPUTED BETWEEN ESTIMATED AND ACTUAL GRADE SCORES FOR EXPERIMENTAL AND CONTROL GROUPS

Source of Estimated/ Actual Correlation Coefficients	Correlation Coefficient	Fisher's Z Transformation	Z-Value	Significance Level
Experimental Groups (N=52)	r = 0.8708	1.3330		
			Z = 6.3334	$p < .0001$
Control Groups (N=50)	C = 0.0401	0.0401		

CHAPTER V

SUMMARY, CONCLUSIONS, AND IMPLICATIONS FOR FURTHER RESEARCH

The purpose of the present study was to determine the effects of self-evaluation training on high school students' ability to estimate the grade the instructor would assign them in a particular course (actual course grade). More specifically, the purpose of the study was to compare the differences between estimated and actual course grades of students who had been taught self-evaluation techniques with the differences between estimated and actual course grades of students who had not been taught self-evaluation techniques. The researcher had hypothesized that those students who had been taught self-evaluation techniques would make more accurate estimates of their course grades than those students who had not been taught self-evaluation techniques.

Students enrolled in four high school English classes were compared on their ability to conduct self-evaluations of their weekly progress and overall class performance. Two classes (Experimental groups; N=52) were trained to conduct self-evaluations of their progress in five course units (Unit Topics). Progress ratings were made on a Student's Self-Evaluation Rating Sheet which had been developed by the student participants, the researcher, and the Chairman of the English Department. Classes three and four (Control

groups; N=50) were not given training in self-evaluation techniques, but were asked to approximate the final grade they would receive in the class. Data collected from these one-hundred two (N=102) students enrolled in high school English courses were used to test six (6) null hypotheses.

It was hypothesized that the self-evaluation training would enable the two experimental groups to make significantly more accurate estimates of the course grades they would actually receive from course instructors than the two control groups. The results of testing the null hypotheses may be summarized as follows:

(1) There was no significant difference between the grades assigned by the course instructor and the grades estimated by the high school students who had been taught self-evaluation techniques.

(2) There was a significant difference between the grades assigned by the course instructor and the grades estimated by the high school students who had not been taught self-evaluation techniques.

(3) The number and magnitude of errors made by the experimental groups (those who had been taught self-evaluation techniques) in estimating their course grades were significantly

less than the number and magnitude of errors made by the control groups (those who had not been taught self-evaluation techniques) in estimating their course grades.

(4) There was a significant relationship between the course grades estimated by the students in the experimental groups and the course grades assigned by the course instructor.

(5) There was no significant relationship between the course grades estimated by the students in the control groups and the course grades assigned by the course instructor.

(6) The relationship between the grades estimated by the experimental group and the grades assigned them by the course instructor was significantly greater than the relationship between the grades estimated by the control groups and the grades assigned them by the course instructor.

Conclusions Made From the Study

The results of testing the hypotheses led to several conclusions. Each of these conclusions is stated below:

The first general conclusion drawn from the results of the study was that teaching self-evaluation techniques

to the high school students in the experimental groups enabled them to make accurate approximations of the grades they would be assigned by the course instructor.

It was further concluded that the students' ability to estimate their course grades was significantly improved after they had been taught self-evaluation techniques. The comparisons made between the grade estimates made by self-evaluating students and estimates made by students who had not been taught self-evaluation techniques further substantiated the first conclusion.

The third general conclusion drawn from the results of the study was that the self-evaluation techniques proved to be a very beneficial teaching tool for the course instructor and a good learning technique for the students. Students in the experimental classes indicated that the evaluation form not only served as a study guide for the English classes, it could be used as a set of learning goals and objectives for each student to accomplish. Some students in the self-evaluating classes further indicated that they had used the evaluation form to make similar estimations of their progress in other classes. With slight alterations, the evaluation sheet can be adapted to any academic area.

The conclusions drawn from the results of the study can be summarized by saying that student self-evaluations can be beneficial to both the student and the instructor if students are trained in self-evaluation techniques.

Implications for Further Research

The implications for further research contained in this section of the dissertation can be categorized into four (4) basic groups: (1) studies which are basically the same as the present study but with variations in the student participants, (2) studies which are basically the same as the present study but with variations in the methods of measurement, data collection instruments, or both, (3) studies which are basically the same as the present study but with variations in the research design, and (4) studies which are basically the same as the present study but with many of the problems eliminated. Specific implications for further research studies are as follows:

Further studies could be conducted in the area of student self-evaluation which would be similar to the present study but with variations in the grade levels of student participants. For instance, student groups could include classes from four different grade levels such as the third, sixth, ninth, and twelfth grades. Such a study would give some indication of the feasibility of student self-evaluation techniques in the lower elementary as well as the secondary grades.

Further research studies could also be conducted which would be similar to the present research effort but the researcher would use different data collection instruments. The data collection instrument used in the present study,

Student's Self-Evaluation Rating Sheet (Appendix B) was developed in cooperation with the course instructor, students and the English Department Chairman. While the areas contained on the Rating Sheet were comprehensive enough to cover the unit topics taught, some assignments were covered more fully than others. This situation could be avoided by further development of the rating sheet. A more comprehensive rating instrument might contain specific questions about the students' competency in the following areas: (1) Understanding of selections, (2) Class participation, (3) Essay Writing and/or lesson preparation, (4) Completion of work assignments, and (5) Extra work performed. A data collection instrument developed in this way would be more comprehensive and yield more information than the rating sheet used in the present study.

BIBLIOGRAPHY

Books

Ahmann, Stanley and Marvin D. Glock, Evaluating Pupil Growth. Boston: Allyn and Bacon, 1967.

Borg, Walter R. Educational Research: An Introduction. New York: David McKay Company, Inc., 1963.

Good, Carter V. Essentials of Educational Research, Methodology and Design. New York: Appleton-Century-Crofts, 1960.

Edwards, Allen L. Statistical Analysis. New York: Rinehart and Company, Inc., 1953.

Faunce, Roland C. and Carroll L. Munshaw. Teaching and Learning in Secondary Schools. Belmont, California: Wadsworth Publishing Company, Inc., 1964.

Grambs, Jean D., John C. Carr and Robert M. Fitch. Modern Methods in Secondary Education. New York: Holt Rinehart and Winston, Inc., 1970.

Hoel, Paul. Elementary Statistics. New York: John Wiley and Sons, Inc., 1966.

Holton, Samuel M. Understanding the American Public High School. Boston: Allyn and Bacon, Inc., 1969.

Hoover, Kenneth. Learning and Teaching in the Secondary School. Boston: Allyn and Bacon, Inc., 1967.

Kerlinger, Fred N. Foundations of Behavioral Research. New York: Holt Rinehart and Winston, Inc., 1967.

Lavin, David E. The Prediction of Academic Performance. New York: Russell Sage Foundation, 1965.

Linquist, E. F. Statistical Analysis in Educational Research. New York: Houghton Mifflin Company, 1940.

Mallery, David. High School Students Speak Out. New York: Harper and Row Publishers, 1962.

Sawin, E. A. Evaluation and the Work of the Teacher. Belmont, California: Wadsworth Publishing Company, Inc., 1969.

Schwartz, Alfred and Stuart O. Tiedman. Evaluating Student Progress in the Secondary School. New York: Longmans, Green and Company, 1957.

Tylor, Ralph W. Educational Measurement. Washington, D. C. American Council on Education, 1951.

Weinberg, George H. and John A. Schumaker. Statistics, An Intuitive Approach. Belmont, California: Wadsworth Publishing Company, Inc., 1967.

Wellington, C. Burleigh and Jean Wellington. Teaching for Critical Thinking. New York: McGraw-Hill Book Company, Inc., 1960.

Wiersman, William. Research Methods in Education. New York: J. B. Lippincott Company, 1969.

Wrinkle, William B. Improving Marking and Reporting Practices. New York: Renshart and Company, Inc., 1947.

Periodicals

Allen, Paul M., "The Student Evaluating Dilemma," Today's Education, LVIII, (February, 1969), p. 48-50.

Barnes, Katherine F. and H. H. Barnes, "A Realistic Approach to Grading," The Clearing House," XXXVI, (April, 1962), p. 476-478.

Coffman, William E. and Dana Kurfman, "A Comparison of Two Methods of Reading Essay Reminations," American Educational Research Journal, V, (January, 1960), p. 99-105.

McGuire, Brian Patrick, "Opinions Differ: The Grading Game," Today's Education, LVIII, (March, 1969), p. 32-34.

Nishsus, Stanley W., "The Anatomy of Evaluation," The Clearing House, XLII, (February, 1968), p. 332-336.

Philbrick, Joseph L. and Patrick L. O'Donnell, "Precision in Grading Practices - Panacea or Problem?" The Journal of Educational Research, XLII, (December, 1968), p. 173-176.

Ransom, Harry H., Ude H. Jansen and James B. Candle, "Educators Look at Grading," The Texas Outlook, XLV, (July, 1961), p. 14-15.

Yauch, Wilbur A., "School Marks and Their Reporting," National Educational Association Journal, L, (May, 1961), p. 50-52.

Encyclopedias

Hagan, E. P. and R. L. Thorndike, "Evaluation," in Encyclopedia of Educational Research, third ed. New York: Macmillan Company, 1960, p. 445.

Dissertations

Crowley, Leon V., "An Investigation of the Marking Practices in Secondary Schools," Unpublished Doctoral dissertation, Department of Education, University of Oklahoma, Norman, Oklahoma, 1968.

APPENDIX A

OBJECTIVES STATED FOR LITERATURE CLASSES

The English Class

Objectives in Literature

Eleventh Grade

General objectives:

1. To find the central meaning of a piece of literature or to understand the plain sense of what has been read.
2. To enlarge the student's oral speaking abilities.
3. To reinforce the student's ability to organize his thinking before discussion.
4. To improve written expression through paragraph writing.
5. To increase a student's knowledge of word meanings.

Specific objectives: (for the "pioneering" unit of literature)

1. To discover the chief themes of American literature
2. To study the nature of American themes and their effect on individuals
3. To examine the role played by the writer in a study of the past
4. To explore the influence of the frontier on American values
5. To deepen the student's appreciation of their American heritage.
6. To document the interpretation of the literature from the literature itself
7. To relate the literature to an individual student's own life
8. To study the nature and uses of setting, characterization, plot, tone, and theme to clarify and intensify a work
9. To develop vocabulary

APPENDIX B

STUDENT SELF-EVALUATION RATING SHEET

STUDENT'S SELF-EVALUATION RATING SHEET

Name: _____ Unit Topic: _____

Period: _____ Date: _____ _____

Directions: Please evaluate your weekly progress in the Unit Topic just completed. Read each of the questionnaire statements carefully. Enter the number which best indicates the weekly progress you have made in studying the Unit Topic listed above. (Use the number codes provided in the box.) In determining your progress, take into consideration your class participation, interests, ability, suggested ideas, cooperation, and attitudes.

The teacher will hold individual conferences with each student to discuss the progress ratings made on this evaluation sheet as well as your progress in other related areas such as grammar usage, writing, spelling, group discussions, etc.

Explanation of Numerical Codes . . .

4 - Superior Progress
3 - Good Progress
2 - Average Progress
1 - Below Average Progress
0 - No Progress

Questionnaire Statement Being Rated	Progress Ratings					Averages
	1st week	2nd week	3rd week	4th week	5th week	
1. I understand the plain sense of the selection I read						
2. I understand the author's central meaning of the selection						
3. I understand what influenced the writer to say what he did						
4. I understand the author's medium (essay, fiction, poetry, etc.) . .						
5. I understand the literary techniques (setting, characterization, conflict, climax, etc.) used						
6. I understand the relevance of the selection to my own life						
7. I took notes in class during the discussion of the unit						
8. I raised questions regarding the material under discussion						
9. I participated in class discussion						
10. I listened attentively to others						
11. I contributed resource material from outside sources (books, periodicals, newspapers, etc.)						
12. I have read additional books to reinforce my knowledge of literature .						
13. I studied for the test covering the unit topic						
14. I understand the mistakes I made on the test over the unit						
15. In essay writing, I followed the proper structure for writing a paragraph (introduction, discussion and conclusion)						
16. I investigated the topic assigned and adapted the materials to the essay						
17. I rewrote my essay to improve my writing ability						
18. I attempted to use the best language I know						
19. I attempted to make my essay clear in meaning						
20. All my sentences contained only one basic statement						
21. I varied the structure of the sentence						
22. The paragraphs developed one basic idea						
23. I understand the grammatical errors I made and corrected them to improve my writing ability						
24. I completed all work assigned in the unit						

APPENDIX C

RAW SCORES AND DESCRIPTIVE STATISTICS OF THE FOUR STUDENT GROUPS

TABLE 8

RAW DATA AND DESCRIPTIVE STATISTICS OF STUDENT GROUP NUMBER ONE (EXPERIMENTAL GROUP ONE)

Student Number	Average Estimated Grade	Actual Class Grade	Estimated/Actual Grade Difference
01	1.00	1.00	0.00
02	1.39	2.00	0.61
03	1.83	1.00	-0.83
04	3.44	2.00	-1.44
05	1.50	2.00	0.50
06	2.46	2.00	-0.46
07	2.66	2.00	-0.66
08	2.00	2.00	0.00
09	3.46	3.00	-0.46
10	2.33	2.00	-0.33
11	1.29	1.00	-0.29
12	2.88	2.00	-0.88
13	1.92	2.00	0.08
14	2.67	3.00	0.33
15	3.37	4.00	0.63
16	2.92	3.00	0.08
17	2.25	2.00	-0.25
18	1.21	1.00	-0.21
19	1.58	1.00	-0.58
20	3.83	4.00	0.17
21	3.48	3.00	-0.48
22	1.08	1.00	-0.08
23	2.08	1.00	-1.08
24	0.88	1.00	0.12
25	1.42	1.00	-0.42
26	1.25	1.00	-0.25
27	3.04	3.00	-0.04
28	2.88	3.00	0.12
29	3.87	4.00	0.13
30	2.96	4.00	1.04
31	0.00	0.00	0.00
32	0.00	0.00	0.00
Mean	2.154	2.000	-0.154
Standard Deviation	1.047	1.136	0.515

TABLE 9

RAW DATA AND DESCRIPTIVE STATISTICS OF STUDENT GROUP NUMBER TWO (EXPERIMENTAL GROUP TWO)

Student Number	Average Estimated Grade	Actual Class Grade	Estimated/Actual Grade Difference
01	3.25	3.00	-0.25
02	2.88	3.00	0.12
03	3.04	2.00	-1.04
04	2.00	2.00	0.00
05	3.21	3.00	-0.21
06	2.67	3.00	0.33
07	2.58	4.00	1.42
08	2.92	4.00	1.08
09	3.50	4.00	0.50
10	2.93	3.00	0.07
11	2.63	3.00	0.37
12	2.74	2.00	-0.74
13	2.71	3.00	0.29
14	2.75	3.00	0.25
15	2.92	3.00	0.08
16	3.46	4.00	0.54
17	1.79	2.00	0.21
18	2.83	3.00	0.17
19	1.58	2.00	0.42
20	3.92	4.00	0.08
Mean	2.816	3.000	0.185
Standard Deviation	0.558	0.725	0.536

TABLE 10

RAW DATA AND DESCRIPTIVE STATISTICS OF STUDENT GROUP NUMBER THREE (CONTROL GROUP ONE)

Student Number	Estimated Grade	Actual Class Grade	Estimated/Actual Grade Difference
01	2.00	2.00	0.00
02	3.00	3.00	0.00
03	4.00	4.00	0.00
04	4.00	4.00	0.00
05	1.00	1.00	0.00
06	4.00	4.00	0.00
07	3.00	3.00	0.00
08	3.00	3.00	0.00
09	2.00	2.00	0.00
10	3.00	3.00	0.00
11	2.00	2.00	0.00
12	3.00	3.00	0.00
13	2.00	2.00	0.00
14	2.00	2.00	0.00
15	2.00	2.00	0.00
16	2.00	2.00	0.00
17	2.00	2.00	0.00
18	0.00	1.00	1.00
19	0.00	0.00	0.00
20	2.00	1.00	-1.00
21	3.00	2.00	-1.00
22	2.00	1.00	-1.00
23	2.00	1.00	-1.00
24	2.00	1.00	-1.00
25	2.00	1.00	-1.00
Mean	2.280	2.080	-0.200
Standard Deviation	1.021	1.077	0.500

TABLE 11

RAW DATA AND DESCRIPTIVE STATISTICS OF STUDENT GROUP NUMBER FOUR (CONTROL GROUP TWO)

Student Number	Estimated Grade	Actual Class Grade	Estimated/Actual Grade Difference
01	2.00	1.00	-1.00
02	3.00	1.00	-2.00
03	4.00	3.00	-1.00
04	2.00	0.00	-2.00
05	1.00	0.00	-1.00
06	2.00	3.00	1.00
07	3.00	2.00	-1.00
08	3.00	2.00	-1.00
09	3.00	2.00	-1.00
10	3.00	2.00	-1.00
11	3.00	3.00	0.00
12	3.00	2.00	-1.00
13	3.00	2.00	-1.00
14	4.00	1.00	-3.00
15	1.00	0.00	-1.00
16	0.00	1.00	1.00
17	1.00	0.00	-1.00
18	3.00	3.00	0.00
19	2.00	3.00	1.00
20	3.00	3.00	0.00
21	3.00	3.00	0.00
22	1.00	0.00	-1.00
23	2.00	2.00	0.00
24	2.00	2.00	0.00
25	2.00	1.00	-1.00
Mean	2.360	1.680	-0.720
Standard Deviation	0.995	1.108	0.936