THE EFFECTS OF WRITTEN COMMENT ON
EXPOSITORY COMPOSITION

APPROVED:

Graduate Committee:

John Greenwell
Major Professor

William McCarter
Minor Professor

Committee Member

Committee Member

Dean of the College of Education

Dean of the Graduate School
THE EFFECTS OF WRITTEN COMMENT ON EXPOSITORY COMPOSITION

DISSERTATION

Presented to the Graduate Council of the North Texas State University in Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

By

Thomas Carroll Gee, B.A., M.Ed.

Denton, Texas

June, 1970
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>v</td>
</tr>
</tbody>
</table>

## Chapter

### I. INTRODUCTION  1

- Statement of the Problem
- Purpose of the Study
- Hypotheses
- Significance of the Study
- Definition of Terms
- Assumptions
- Limitations of the Study
- Procedures for Collecting Data
- Procedures for the Treatment of Data

### II. RELATED LITERATURE  14

- Conditioning in the Classroom
- Positive Reinforcement
- Negative Reinforcement
- No Reinforcement
- Feedback---Vehicle for Reinforcement
- Effects of Reinforcement
- Learning Theory and Written Composition

### III. PROCEDURES FOR OBTAINING AND TREATING DATA  36

- Description of the Subjects
- Procedures for Collecting Data
- Analysis of the Data
- Statistical Treatment of the Data

### IV. PRESENTATION AND ANALYSIS OF DATA  43

- Findings for Total Subjects
- Findings for High-Ability Subjects
- Findings for Middle-Ability Subjects
- Findings for Low-Ability Subjects
- Summary of Findings
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Distribution of Subjects by Ability</td>
<td>37</td>
</tr>
<tr>
<td>II. Examples of Comments for Each Ability Group</td>
<td>38</td>
</tr>
<tr>
<td>III. Mean Gains and Standard Deviations for Quantitative Measures for Total Subjects</td>
<td>44</td>
</tr>
<tr>
<td>IV. Summary of Analysis of Variance of Number of T-Units for Total Subjects Using Three Types of Criticism</td>
<td>44</td>
</tr>
<tr>
<td>V. Duncan's Multiple Range Test of Significance Between Pairs of Means for Number of T-Units for Total Subjects</td>
<td>45</td>
</tr>
<tr>
<td>VI. Summary of Analysis of Variance of Quality Ratings for Total Subjects Using Three Types of Criticism</td>
<td>46</td>
</tr>
<tr>
<td>VII. Mean Differences and Standard Deviations for Quantitative Measures for High-Ability Subjects</td>
<td>47</td>
</tr>
<tr>
<td>VIII. Summary of Analysis of Variance of Number of T-Units for High-Ability Subjects Using Three Types of Criticism</td>
<td>47</td>
</tr>
<tr>
<td>IX. Duncan's Multiple Range Test of Significance Between Pairs of Means for Number of T-Units for High-Ability Subjects</td>
<td>48</td>
</tr>
<tr>
<td>X. Summary of Analysis of Variance of Quality Ratings for High-Ability Subjects Using Three Types of Criticism</td>
<td>49</td>
</tr>
<tr>
<td>XI. Mean Differences and Standard Deviations for Quantitative Measures for Middle-Ability Subjects</td>
<td>50</td>
</tr>
<tr>
<td>Table</td>
<td>Summary of Analysis of Variance of</td>
</tr>
<tr>
<td>------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>XII.</td>
<td>Number of T-Units for Middle-Ability Subjects Using Three Types of Criticism</td>
</tr>
<tr>
<td>XIII.</td>
<td>Summary of Analysis of Variance of Quality Ratings for Middle-Ability Subjects Using Three Types of Criticism</td>
</tr>
<tr>
<td>XIV.</td>
<td>Mean Differences and Standard Deviations for Quantitative Measures for Low-Ability-Level Subjects</td>
</tr>
<tr>
<td>XV.</td>
<td>Summary of Analysis of Variance of Number of T-Units for Low-Ability-Level Subjects Using Three Types of Criticism</td>
</tr>
<tr>
<td>XVI.</td>
<td>Summary of Analysis of Variance of Quality Ratings for Low-Ability Subjects Using Three Types of Criticism</td>
</tr>
<tr>
<td>XVII.</td>
<td>Summary of t-Test for Praised Males and Females on Quantified Data</td>
</tr>
<tr>
<td>XVIII.</td>
<td>Summary of t-Test for Criticized Males and Females on Quantified Data</td>
</tr>
<tr>
<td>XIX.</td>
<td>Summary of t-Test for No Comment Males and Females on Quantified Data</td>
</tr>
<tr>
<td>XX</td>
<td>Summary of Chi Square Test for Difference Between Attitudes of Males and Females Who Were Praised</td>
</tr>
<tr>
<td>XXI.</td>
<td>Summary of Chi Square Test for Difference Between Attitudes of Males and Females Who Were Criticized</td>
</tr>
<tr>
<td>XXII.</td>
<td>Summary of Chi Square Test for Difference Between Attitudes of Males and Females Who Received No Comment</td>
</tr>
<tr>
<td>XXIII.</td>
<td>Summary of Chi Square Test for Differences in Attitudes Between Those Who Were Praised and Those Who Were Criticized</td>
</tr>
<tr>
<td>XXIV.</td>
<td>Summary of Chi Square Test for Differences in Attitudes Between Those Who Were Criticized and Those Who Received No Comment</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>XXV.</td>
<td>Summary of Chi Square Test for Differences in Attitudes Between Those Who Were Praised and Those Who Received No Comment</td>
</tr>
<tr>
<td>XXVI.</td>
<td>Summary of Findings for Quantitative Data for Three Treatment Groups</td>
</tr>
<tr>
<td>XXVII.</td>
<td>Summary of Findings for Attitudes of Three Treatment Groups</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

One of the most difficult tasks of English teachers, and perhaps any teacher who must evaluate a written essay, is in deciding just how it should be graded. In 1965 a study of English by the National Council of Teachers of English (5) and a barrage of independent studies such as Dusel's (4) began to question the validity of teachers' approaches to the teaching of composition. Some suggested that the frustrations of evaluating essays came as a result of poor training in writing and in evaluating essays. From the 1965 N.C.T.E. report came the realization that English teachers, nation-wide, had only sixteen to eighteen hours of collegiate study in English and that fewer than half of the English teachers across the nation had majored in English (5, p. 5). Muller summarizes some of the confusion in teaching composition by saying, "English teachers go on as if they agreed with the general public that composition is absolutely essential, much more important and practical than literature; yet nothing in English is taught less effectively, amid more confusion and conflict of theory, or hunch" (11, p. 98).

In the past decade practically every aspect of teaching composition has been questioned, from who should teach it to
how much it should be taught, from how composition should be evaluated to what methods are most effective in teaching. One of the specific concerns teachers have begun to investigate is the effect of teacher comments on subsequent student achievement. Both psychologists and English teachers have become more aware of the psychological effect that their comments may have; but, again, methods are too often guided by hunch (7, p. 16; 8, p. 986; 9, p. 118). The next step should be to ask further questions, to put common-sense hunches to test.

Statement of the Problem

The problem of this study was to determine the effects of praise, criticism, and no response on the expository compositions of eleventh-grade students, as measured by three raters for quality, by number of T-units for quantity, and by an attitude questionnaire.

Purpose of the Study

This study was planned to investigate the effects of written comments on the expository compositions of eleventh-grade students using a cross section of ability groups. Data for combined groups and data for high-, middle-, and low-ability groups were used to determine whether one type of comment was more effective than another in improving the quantity and quality of student compositions and in improving student attitudes toward composition. Teachers may use the
findings as a guide to what kinds of comments are most effective in reinforcing good writing skills and attitudes.

Hypotheses

Solution to the problem was sought by testing these hypotheses:

I. Subjects in the following levels who are praised will show significantly greater mean gains in quantity of their compositions than will subjects who receive criticism:
   A. All levels combined,
   B. High-ability-level subjects,
   C. Middle-ability-level subjects,
   D. Low-ability-level subjects.

II. Subjects in the following groups who are criticized will show significantly greater mean gain in quantity of their compositions than will subjects who receive no comment:
   A. All levels combined,
   B. High-ability-level subjects,
   C. Middle-ability-level subjects,
   D. Low-ability-level subjects.

III. Subjects in the following levels who are praised will show significantly greater mean gain in quality of their compositions than will subjects who receive criticism:
   A. All levels combined,
   B. High-ability-level subjects,
   C. Middle-ability-level subjects,
   D. Low-ability-level subjects.
IV. Subjects in the following groups who are criticized will show significantly greater mean gain in quality of their compositions than will students who receive no comment:

A. All levels combined,
B. High-ability-level subjects,
C. Middle-ability-level subjects,
D. Low-ability-level subjects.

V. Within the treatment group receiving praise there will be no significant difference between sexes on measures of

A. Mean gain in quantity,
B. Mean gain in quality,
C. Attitudes.

VI. Within the treatment group receiving criticism there will be no significant difference between sexes on measures of

A. Mean gain in quantity,
B. Mean gain in quality,
C. Attitudes.

VII. Within the treatment group receiving no comment there will be no significant difference between sexes on measures of:

A. Mean gain in quantity,
B. Mean gain in quality,
C. Attitudes.
VIII. Subjects who are praised will show attitudes which are significantly more positive than those who receive criticism.

IX. Subjects who are criticized will show attitudes which are significantly more positive than those who receive no comment.

Significance of the Study

An investigation of the effects of teacher comments on student-written compositions is significant for several reasons. First, according to basic tenets of learning theory, a response which is praised tends to be repeated. Travers concludes that the study of verbal behavior "constitutes an important educational activity" and that reinforcement of verbal behavior may have important educational objectives (15, p. 118). On the other hand, if a student receives no comment, extinction will follow. While many experiments have tested the reinforcement-non-reinforcement hypothesis, Sweet reports that "very few investigations have studied the effects of written comments on test papers" (13, p. 1).

Present research in the effects of written comment on written composition is inconclusive and even contradictory. While the theoretical assumptions about reinforcement have been tested in the laboratory, they have not been adequately tested in the high school classroom. While some studies of reinforcement have been made in art and creative writing, no
thorough study has been made on expository writing. In none of the current literature has there been a report of a careful study on the effects of reinforcement on ability levels. It also becomes apparent in reading current studies that in judging quality of compositions and reinforcement, previous studies have evaluated "quality" on limited criteria rather than on a realistically broad range of writing skills. What is needed, then, is a carefully controlled study to determine the effects of teacher comments on written composition.

Definition of Terms

For the purpose of this study, the following definitions were established:

1. **Expository writing** refers to writing that is meant to explain. Within the body of the composition there may be argumentation, narration, and description, but the over-all composition will have explanation as its purpose.

2. **T-Unit** is a minimal terminable unit made of one main clause plus the subordinate clauses attached to it or embedded within it (7, p. 20).

3. **High-ability group** is composed of those subjects who scored between 141 and 114 on the California Short-Form Test of Mental Maturity, S Form, Level 4.

4. **Middle-ability group** is composed of those subjects who scored between 113 and 92 on the California Short-Form Test of Mental Maturity, S Form, Level 4.
5. **Low-ability group** is composed of those subjects who scored 91 or below on the *California Short-Form Test of Mental Maturity, S Form, Level 4*.

6. **Comments of praise** were those comments which encouraged the student by calling attention to the best aspects of his writing and ignored the poor aspects of his writing.

7. **Comments of criticism** were those comments which called attention to the poorer aspects of the student's writing and ignored the good aspects of his writing.

8. **No comment** referred to the lack of any written comment on the student's paper.

**Assumptions**

The basic assumptions of this study were

1. That the four writing stimuli had equivalent motivational interest and difficulty.

2. That the subjects cooperated and performed at their best level.

**Limitations of the Study**

1. This study was limited to a consideration of the quantity and quality of writing and to student attitudes toward writing.

2. This study was limited to expository writing.
Procedures for Collecting Data

Sample

Seven sections of junior English students in a north central Texas high school were subjects in this study. Four of the sections were considered average classes by the school. Two were considered basic classes, and one was considered accelerated. From the 150 students in the classes, the 139 who completed all four writing trials constituted the experimental sample. Results on the California Short-Form Test of Mental Maturity were used to divide them into high, middle, and low ability groups.

Procedure

During the fall of 1969, appropriate preliminary measures were taken. Permission to conduct the study was requested from the superintendent of the school system and from the principal and English teachers of the high school.

From two teachers' classes, seven sections of eleventh-grade English were selected. Teacher selection was based on willingness to cooperate in the experiment. To minimize possible influences of the teachers on the outcome of the experimental variable, the following precautions were taken:

1. Teachers were recommended by the department chairman as having similar styles of teaching and criticism.

2. Teachers agreed to follow the same curriculum plan during the experiment.
3. Teachers agreed to offer no instruction, discussion, or evaluation of composition during the time of the experiment.

4. The investigator gave instructions for the writing trials, administered the intelligence tests, and returned the compositions to the students.

Four topics were used as stimuli for writing. The topics were selected as equivalent in motivational interest and difficulty by a panel of four English teachers who had experience in teaching junior English.

Fifty minutes were allotted for each writing trial (1, p. 9). Compositions were then collected and marked with teacher-comments appropriate to the group to which the student had been assigned. One week later the student's previous composition was returned to him, and he was asked to re-read his paper and consider the comments that had been made. If he had not received comments, he was asked to re-read his paper looking for ways he could improve his next essay.

Analysis of Data

After the last composition had been written, papers I and IV were analyzed for quality and quantity to determine mean gains. Quantity was determined by computing the number of T-Units. Quality was judged by a panel of three certified English teachers who established a .72 reliability in scoring
before they evaluated the papers in the experiment. Quality was based on the following criteria established by Diedrich (3, p. 24) (see Appendix A):

<table>
<thead>
<tr>
<th>Category</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ideas</td>
<td>1-5</td>
</tr>
<tr>
<td>Form</td>
<td>1-5</td>
</tr>
<tr>
<td>Flavor</td>
<td>1-5</td>
</tr>
<tr>
<td>Mechanics</td>
<td>1-5</td>
</tr>
<tr>
<td>Wording</td>
<td>1-5</td>
</tr>
<tr>
<td>Total</td>
<td>5-25</td>
</tr>
</tbody>
</table>

After each rater independently rated each composition, the ratings of the three judges were averaged. Raters scored only thirty papers at one sitting to avoid a fatigue factor. Students were asked to identify their papers only by a number which had been assigned to them to avoid any possible bias on the rater's part.

The attitudes of students were determined by a questionnaire that was quantified by a yes-no response (see Appendix B). The questionnaire was given after the last composition had been returned to the students.

**Treatment of Data**

A one-way analysis of variance was computed to test Hypotheses I, II, III, and IV. Where significant F's were found, a Duncan's New Multiple Range Test was used to determine which means differed significantly.
Parts A and B of Hypotheses V, VI, and VII were tested by a t test. Part C of Hypotheses V, VI, and VII and Hypotheses VIII and IX were tested by using Chi Square. The .05 level of confidence was required for significance.


13. Sweet, Roger C., Educational Attainment and Attitudes Toward School as a Function of Feedback in the Form of Teachers' Written Comments, Madison, Wisconsin, University of Wisconsin, November, 1966, Educational Resources Information Center (ED 015 163).


CHAPTER II

RELATED LITERATURE

Anyone concerned with learning is obviously faced with questions regarding reinforcement. How can one reinforce a particular behavior? How can one extinguish a particular behavior? What events will cause an individual's behavior to be reinforced?

For the English teacher the questions might be stated more specifically. How can one reinforce the behavior of a student who has written a good composition? How can one best encourage a student to proofread more carefully or to avoid punctuation errors? Will criticism of his errors, praise of his composition skills, or disregard of his faults bring about the most desirable change?

At the base of all of these questions is the concept of reinforcement. Unfortunately, however, agreement is lacking among psychologists about why a particular event, whether positive or negative, reinforces as it does. Some say an event reinforces because it acts as a reward; others say that the event reinforces because it provides information to the subject; still others say that an event acts one way in a particular situation and another way in a different situation. What is important is that behavior can be modified by
reinforcement although we do not fully understand how the reinforcement fostered change.

Using the basic tenets of Stimulus-Response Theories, there have been several major studies completed in the 1960's which have studied the effects of reinforcement (3, 5, 6, 29, 30, 37, 38). The literature pertinent to this research is presented in two sections: (1) background and principles of reinforcement, and (2) studies that test reinforcement theories in the teaching of written composition.

Conditioning in the Classroom

Reinforcement refers to an event which, when following closely upon a behavior, increases the frequency with which the organism engages in that behavior. A positive reinforcer is an event, the presentation of which constitutes a reinforcement. Originally, B. F. Skinner defined a negative reinforcement as an event the removal of which is reinforcing. The central idea in negative reinforcement is that once the organism has learned to anticipate the aversive stimuli [criticism on written papers], he will avoid the stimulus by behaving differently. The actual operations and definitions of negative reinforcement are somewhat complicated by later writings of E. L. Thorndike. Thorndike assumed that praise and punishment had opposite effects, that praise ensured that learning took place, and punishment was effective in unlearning (39, p. 82). Hilgard and Bower have sifted the ideas on
reinforcement from learning theorists prior to 1966 and have concluded that, with few exceptions, all learning theorists could agree that when one talks of reinforcements, positive reinforcements are equivalent to rewards and successes and that negative reinforcements refer to punishments and failures (19, p. 563).

Reinforcing properties are defined in terms of their effects upon behavior. It would be impossible, for example, to interpret a single event as either positively or negatively reinforcing except in terms of its effect on behavior. In other words, what might be satisfying to one student might be annoying to another. However, the most common way teachers have of providing reinforcement is by offering verbal or written comments, which might be called feedback, wherein the student receives information about the nature and effect of his behavior. In practice, teachers seldom offer one or the other but use both positive and negative comments simultaneously in an effort to help students improve.

Positive Reinforcement

Skinner implies that there is no real problem in finding reinforcers for human behavior, that the most common and most effective way in a learning situation is to tell the learner that the response made was the correct one (34). In the classroom teachers reward more than just a perfect
performance. In the process of learning, close approximations of a perfect performance are rewarded with the expectation that in due time the student will come closer and closer to the expected behavior or performance level. The concept of shaping, as its name implies, is the technique by which successive approximations of the expected behavior are reinforced. Travers concluded that positive reinforcement alone can change speech habits (and perhaps writing habits) even though the learner may be unaware of the fact that his behavior is being modified. Travers concluded by saying that teachers need to employ positive reinforcement for effecting proper speech because "it is this practice that in the long run will make for desirable habits" (39, p. 91).

Hilgard and Bower concur that verbal approval is a vehicle of reinforcement by saying that among the most common secondary reinforcers are praise, social approval, attention, and spoken exclamations such as "good" (19, p. 122). Apparently, then, Hilgard and Bower view the teacher's verbal estimate of liking the student's work or finding his work attractive as one of the most valid, general reinforcers for humans.

E. B. Page's study of teacher comments and student performance found that positive reinforcement in the form of encouraging and/or constructive comments did more to increase learning than did criticism which ignored the student's work or gave little directive comment. He concluded
When the average secondary teacher takes the time and trouble to write comments (believed to be encouraging) on student papers, these apparently have a measurable and potent effect upon student effort, or attention, or attitude, or whatever it is which causes learning to improve, and this effect does not appear dependent on school building, school year, or student ability (30, p. 181).

Negative Reinforcement

Where positive reinforcement increases the frequency with which the student engages in a behavior, negative reinforcement, causes the student to avoid an aversive stimuli by learning a new behavioral pattern. Negative reinforcers, as Hively notes, may produce many undesirable side-effects such as timidity, anxiety, and aggression which may conflict with learning (20, p. 39). According to Solomon and Rosenberg's analysis, positive reinforcement indicates approval, negative reinforcement indicates disapproval, and no reinforcement indicates statements which can be interpreted as neither positive nor negative (35, p. 199).

If the student can find a suitable avoidance response for negative reinforcements, there can be effective learning. Too often in composition, however, the student may avoid writing altogether rather than avoiding specific errors. A study by Travers, Reid, and Nelson concluded that a prolonged aversive stimulus might tend to build up an avoidance response that would become generalized to all learning situations (39, p. 84). Guthrie suggests that action leading to pain tends to be stamped out (17, p. 82). To insure
that all desire to learn is not stamped out, teachers must be especially careful in using negative reinforcement. Lazarus warned that a long series of failures reduces both the level of performance and general intellectual functioning (23, p. 297).

While learning theorists do not fully agree on the exact way in which negative reinforcement operates, Hilgard and Bower summarize what agreement does exist. They say, "It is generally found that positive reinforcements (rewards, successes) are to be preferred to negative reinforcements (punishments, failures)" (19, p. 563).

No Reinforcement

Most psychologists would agree that responses can be weakened or extinguished by continued nonreinforcement (21, p. 217). Like negative reinforcement, no reinforcement may have aversive side-effects. Hilgard and Bower point out that when a subject expects a reward and is not given one, primary frustration reactions occur (19, p. 488). Since extinction (elicitation without confirmation) produces repeated frustration at the goal, the subject may come to anticipate frustration. Nonreward, when reward is expected, has an aversive effect much like punishment.

Feedback—Vehicle for Reinforcement

There have been those who see learning as its own reward, those who attribute the learning process to something
inside the individual. Those who subscribe to Guthrie's theory of contiguity believe that when a stimulus and response are frequently connected closely that the stimulus will come to elicit the response in further situations. Teachers often assume that this theory is valid. English teachers, for example, show models of good writing and expect students to imitate the model, or they ask students to pronounce a word assuming that proper pronunciation will be sufficient to effect proper spelling.

Learning theorists, even if they accept the theory that reinforcement can be intrinsically generated, believe that most students are not sufficiently intrinsically motivated, that they require some form of reinforcement from the outside. Feedback is generally defined as information the student receives about the nature of his performance. To provide information teachers often point out relationships of the subject matter to the student's interests and give or withhold attention and approval. The teacher's manipulation of feedback, then, become a vehicle for providing reinforcement.

Students receive feedback in many forms. Teachers go beyond an evaluative comment such as "That's correct" or "You need to improve" by providing information that identifies the exact nature of the correctness or error. The student also receives feedback from his own judgments, from cues and comments of other students in the class. As Mowrer
suggests, "Behavior is a continuous, on-going function of the informational from all senses, internal and external . . ." (27, p. 309).

Without feedback, students do not know if their behavior approximates what is expected. Although they may be intrinsically satisfied, they gain assurance by evaluation from others. E. L. Thorndike, for example, found that when he asked students to draw three-inch lines without a guide that students could not improve without feedback. A study by Travers, Reid, and Nelson also concluded that "learning is dependent upon the amount of information transmitted by the signal . . ." (39, p. 84). A student who receives information about his writing is in a position to alter or continue his present performance level, but without feedback, he has little reason, beyond intrinsic motivation, to change.

'Summary of Effects of Reinforcement

Silberman reports that in observing forty teachers who were characterized as being from supportive to reproving in the classroom, their comments had little relation to pupil growth in reading. Perhaps this finding indicates that the students' intrinsic reinforcement or peer reinforcement was more important to the students than the teachers' comments. Most research, certainly, contradicts Silberman's findings (1, 6, 12, 24, 30, 37, 38). Students seem to need information, reward, or advice following behavioral events. It is
through such feedback that an individual gains reassurance and direction. Some students, of course, need more external reinforcement than others.

Hively emphasizes the importance of external reinforcement and especially the teacher's role in giving feedback. He says

... the apathy resulting from infrequent reinforcement, the misinformation resulting from slipshod contingencies of reinforcement, the misbehavior resulting from conflicting sources of control, and the anxiety, avoidance and resentment resulting from negative reinforcement are all too common in the average classroom (20, p. 40).

Travers (39), Hilgard and Bower (19), Page (30), Mowrer (27), and Solomon and Rosenberg (35) also place importance on the teacher's comments in the learning progress of students. Generally, they agree that (1) positive reinforcement provides encouragement and becomes associated with pleasurable events, (2) negative reinforcement is effective at time, but it is unpredictable and often damaging, producing anxiety or aggression, and (3) to offer no reinforcement will gradually extinguish a response and perhaps even have an aversive effect much like punishment.

Learning Theories and Written Composition

Learning to write is a cumulative and complex process. Just what part of becoming a good writer is dependent on the teacher and her ability to teach good writing techniques is
probably unmeasurable, but psychologists and teachers assume that the role of the teacher is significant. She not only offers models, demonstrations, lessons, and drill, but she spends hundreds of hours writing comments on students' papers in the belief that her comments will bring about a change--a change for the better.

Since the 1960's there have been hundreds of articles and dozens of books written on the teaching of composition. Perhaps the current prolific study of composition is testimony of the concern English teachers have about the teaching of composition. They want their students to become better writers, and in that sense, they want to shape the behavior of their students and to offer the kinds of reinforcements that will make writing as pleasurable as possible and as worthwhile as possible. Teachers want to avoid the kinds of reinforcement that would cause a student to look toward composition with disgust or fear.

Under contract with the United States Office of Education, Robert Sweet tested the effects of feedback on subsequent student attainment and student attitudes toward English (37). He found that free or specified comments had little effect on test performance and that free comments over a longer period of time significantly increased school performance and significantly changed attitudes in a positive direction. "Teachers' comments," he said, "are like letter grades, the comments are a reinforcement component providing
feedback to the student about some of the effects of his behavior" (37, p. 1).

While Sweet's experiment was tightly controlled, he did not clearly measure the effects of written comments alone because papers were returned to students with numerical and letter grades, thus introducing a possible spurious variable, grades, making his results questionable. It is also important to note that in Sweet's experiment students were assigned to a treatment group based on their previous grades in English; therefore, no conclusion regarding differences among ability groups could be drawn because a standardized measure of ability was not a criterion for placement in a treatment group. Sweet also failed to investigate whether reinforcement affected boys and girls differently.

Taylor's study sought to determine the effects of praise upon the production of creative writing in a fifty-grade science class by analyzing quantity and quality of writing and attitudes of the pupils toward creative writing (38). He found no significant differences in the creative writing endeavors of two groups working under varying conditions. He was able to accept the hypothesis that those who received praise did produce a higher quantity of writing than did groups who received no praise. Taylor also concluded that students who received praise had the most favorable attitudes toward the experiment. In addition, Taylor found that those who did not receive praise manifested
overt indications of frustration, hostility, and dissatisfaction. In conclusion, he said, "Praise does not diminish the quality of work done, while it does increase the quantity and tends to make pupils more independent" (38, p. 83). Such findings suggest that it would be advantageous for teachers to reduce criticism of errors and to increase praise of strengths of students' work.

In evaluating quality of composition, Taylor was concerned solely with the creativity of the composition, not with organization, paragraphing, sentence structure, or mechanics. While creativity is, of course, a major concern to teachers, it is only one criterion which teachers normally use in evaluating a student's paper. Taylor failed to measure the effects of reinforcement on a wide range of variables; thus, he left a great deal to be investigated. In his recommendations Taylor called for further analysis of reinforcement of written composition in relation to ability groups.

The studies of Burton and Arnold offer evidence somewhat contradictory to Taylor's findings (5). In a 1963 study they concluded that no significant differences were found to be associated with intensity of evaluation on frequency of writing, but they, too, stressed the need for further research.

In another study, Buxton sought to determine whether students in a free atmosphere who received no grades or marks
but only comments of praise or students in a revision group who got grades, thorough marking of errors, and were required to revise, made the most progress in writing (6).

He found that the revision group made the most significant improvement. His controls, however, are obviously imprecise. It is impossible to isolate from his study whether the grades, the markings, the process of revision, or a combination of the variables made the revision group's progress more significant than that of the free group. Hagestrom points out that Buxton could offer no explanations of the relative effect of these variables (18, p. 54).

In a study to determine the effect of teacher comments on student performance, Page selected 74 secondary teachers and 2,139 students (30). The students were divided into three groups, one group receiving no comment, one group receiving whatever comments the teacher felt appropriate, and one group receiving certain uniform comments designated by the experimenter that were thought to be generally encouraging. Page's study was carried out to measure the effects of teacher comment on subsequent objective tests. He found that those who received comments that the teacher felt appropriate did best on subsequent tests and that those who received specified comments did better than the no comment group. Page's study provides evidence that a combination of both information (grades) and reinforcers (personalized
comments) is superior to informational feedback alone in positively affecting scholastic performance.

The above studies have called attention to the success of praise in encouraging written composition based on increased quantity and quality and more positive attitudes. There are also numerous studies, based on experience and opinion rather than on empirical research, which support the idea that praise is desirable in fostering good attitudes toward writing (1, 6, 7, 11, 13, 22, 28, 29, 41). Gage's Handbook of Research on Teaching reports that little research has been done on the psychological response and attitude of teacher comments on student writing (25, p. 986). Mirrieless believes that English teachers discourage the good brains of our schools by absurd evaluations (26, p. 118).

Dusel found that teacher comments were not only ineffective but frequently harmful. He warned that "young persons attributed varied meanings to terms frequently applied to their writing, such as 'awkward,' 'weak,' 'disorganized,' 'confused,' and 'incoherent!" (11, p. 986). Dusel reminds teachers that they should be more understanding and tactful than students, more aware of the power of fault-finding to discourage and thus to inhibit learning. Essentially, Dusel repeats the warnings of the learning theorists that negative reinforcement or no reinforcement at all may foster anxiety, apathy, or even fear.
McColly feels that increasing maturity or instructional factors, rather than teacher corrections, annotations, and ratings will remedy most of the problems of writing weakness (24, p. 16). Dendy adds support to this theory (7). In a study to determine whether or not students whose teachers possess "more strict" standards in composition instruction perform differently on a standardized essay test than do those whose teachers have "less strict" standards found that there were no statistically significant differences in writing performance.

Fisher is a strong advocate of removing the teacher from the role of evaluator. He suggests that students want response but that they do not necessarily want criticism. In writing, as in art, students seek teacher approval as well as personal satisfaction. Warmth, support, and appreciation will generally free the child for further exploration, he says (12, p. 209). Fisher, then, is following the basic idea of Hilgard and Bower that praise is the best way of insuring the repetition of a behavior. He says

Creative writing can be one of the most satisfying language experiences for elementary school children. It is very possible that we have stifled much creativity by trying to fulfill too many purposes simultaneously. Art educators have developed procedures to release creative potential in young children, and they do it largely by nondirective procedures relying heavily upon praise in alliance with the developmental needs of children (12, p. 211).
Applegate's advice to English teachers is like Fisher's. She suggests that English teachers become more like art teachers, helping each child to improve in his writing power and clarity but doing it through suggestion and praise and further teaching. As she says, "Comments on creative writing must be appreciative; they must be positive; they must give children goals to work for . . ." (1, p. 54). In fostering a good classroom atmosphere, teachers must create a place without fear.

While many studies have focused on the importance of providing positive reinforcement, others have concentrated on the possible harmful effects of negative reinforcement. Burrows, for example, reminds teachers that fear of failure is an important cause of infrequent writing (4, p. 91). The National Council of Teachers of English also reminded teachers that the teacher "must not kill whatever it is that makes a student willing to try his hand but owes it to him also to show him that when the authors talk of the 'lonely misery' of writing, they are not talking about first drafts" (14, p. 106). Applegate adds that writing is an art but that teachers have taught it as if it were an exact science like arithmetic. "Art," she says, "is to be appreciated, but writing to be criticized. At least that has been our practice if not our theory" (1, p. 33). Muller says that too often we try to shame students into becoming good writers (28, p. 125). He suggests that if teachers would be more
sympathetic and allow students to develop a sense of enjoyment from writing that they would naturally improve simply because they naturally want to improve what they are interested in.

Although not immediately concerned with specific kinds of reinforcement, several writers have emphasized the need for teachers to give some form of feedback. Warriner, in developing a composition text for high school students, suggested to teachers that "An unread, uncorrected, unreturned theme was not worth assigning" (40, p. 17). Following the theories of reinforcement, the teacher could expect the student to become disgusted with writing. Bilodeau and Bilodeau stated, "Studies of feedback or knowledge of results show it to be the strongest most important variable controlling performance and learning" (2, p. 280). Green agrees that the terminal comment is an important instructional device. He says its function should be to identify the major weaknesses of the paper and to reveal to the student ways in which he can improve his composition (16, p. 215). Green views the comments as instructional feedback which can reinforce the student's commendable qualities in writing and suggest ways of improving other qualities.

Burrows says, "We must build by praise, but it must be discriminating. Say something positive and constructive, or nothing" (4, p. viii). Schvaneveldt, in an extremely thorough review of the informational component of feedback
stated that a performance-related signal may be called anything from reward to knowledge of performance because of the dimensions on which it could vary (32).

Summary

Learning theorists and many English teachers are beginning to place a great deal of importance on the role of teacher comments on student compositions. Although research is still limited, several major experiments have been underway during the 1960's to test the effects of teacher comments on subsequent student achievement in composition. The kinds of studies that have been underway reflect the concern English teachers have for the teaching of composition, and it will likely be through studies of isolated variables such as the effect of feedback, models, frequency of writing, etc. that some answers to the question of how best to improve student writing will come.

Both learning theorists and English teachers believe that the teacher's feedback to students may be one of the most potent sources of reinforcement for students. Through the teacher's comments students may be encouraged to improve their writing or to resent writing altogether. Such beliefs are not without support, but classroom studies of reinforcement to date are not conclusive.
CHAPTER BIBLIOGRAPHY


5. Burton, Dwight L. and Lois V. Arnold, Effects of Frequency of Writing and Intensity of Teacher Evaluation Upon High School Student's Performance in Written Composition, Florida State University, Tallahassee, 1963, Educational Resources Information Center (ED 003 281).


37. Sweet, Robert C., Educational Attainment and Attitudes Toward School as a Function of Feedback in the Form of Teachers' Written Comments, University of Wisconsin, Madison, Wisconsin, 1966, Educational Resources Information Center (ED 015 163).


CHAPTER III

PROCEDURES FOR OBTAINING AND TREATING DATA

This chapter is a detailed description of the subjects, the procedures used for collecting data, and the methods employed in analyzing data and the statistical treatment of the data.

Description of the Subjects

Students in seven eleventh-grade English classes in a north central Texas city were subjects in this study. All of the classes were in the same senior high school. Five of the sections were taught by one teacher, two sections by another teacher.

The California Short-Form Test of Mental Maturity, S Form, Level 4 was administered to the classes by the experimenter one week prior to the collection of the other data. The tests were machine scored and interpreted according to the test manual (3). The total intelligence quotient was used to divide the subjects into high, middle, and low ability groups as shown in Table I. High ability was defined as I.Q.'s between 141 and 114, middle ability as I.Q.'s between 113 and 92, and low ability as I.Q.'s 91 or less.
Any student not present for the four writing trials was eliminated from the study. Complete data were compiled for 139 of the original 150 subjects.

Procedures for Collecting Data

Data were collected over a four-week period. Classroom teachers helped select the topics for the writing trials and were able to choose four topics which they thought were equivalent in motivational interest and difficulty. Topics which students had discussed in class or had written compositions on during the year were eliminated (see Appendix C). No instruction in composition or discussion of the compositions in the experiment was made by the teachers during the time of the experiment. Students were given fifty minutes to write on each topic.

Before the topics were given to the students, standardized instructions regarding the experiment were read to the students (see Appendix D). After the first composition had been written, the experimenter read all papers and wrote
comments appropriate to the treatment group to which the student was assigned. Table II describes the kinds of comments that were included for the praised and criticized groups. Those in the no comment group were given only a check mark showing that their papers had been read.

### Table II: Examples of Comments for Each Treatment Group

<table>
<thead>
<tr>
<th>Criticized Group</th>
<th>Praised Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. This doesn't make sense.</td>
<td>1. You raise some interesting questions.</td>
</tr>
<tr>
<td>2. You did not follow the topic assigned.</td>
<td>2. Cleverly said!</td>
</tr>
<tr>
<td>3. This is awkward and unclear.</td>
<td>3. Your grammar and sentence structure are much better in this paper.</td>
</tr>
<tr>
<td>4. Use your imagination. Most of your ideas are trite.</td>
<td>4. This paragraph is a vivid description.</td>
</tr>
<tr>
<td>5. Please proofread; your grammar and spelling are poor.</td>
<td>5. This paragraph establishes good unity for your essay.</td>
</tr>
<tr>
<td>6. Choose fresh, effective words instead of these.</td>
<td>6. This is interesting. Keep up the good work.</td>
</tr>
<tr>
<td>7. Stringy sentences, illogically connected.</td>
<td>7. A very original approach to the topic!</td>
</tr>
<tr>
<td>8. You contradict this idea in your next paragraph.</td>
<td>8. These ideas show a keen insight into the problem.</td>
</tr>
<tr>
<td>9. You ramble. Please organize more carefully.</td>
<td>9. The precise vocabulary that you use here makes your writing fresh and enjoyable.</td>
</tr>
<tr>
<td>10. You over-simplify the problem.</td>
<td>10. I hope your grammar and usage are always this good.</td>
</tr>
</tbody>
</table>
The criticized group's papers were marked for specific errors in grammar, spelling, organization, usage, and suggestions were made in content and style. Only errors and suggestions for improvement were made; the good aspects of the writing were ignored.

The praised group's papers were complimented for their good aspects. Originality, sound and thoroughly developed ideas, good grammar, etc. were praised. Errors were ignored, and no suggestions for improvement were made. No grade was marked on any of the papers.

One week from the time the first papers were written, they were returned to the students. Students were asked to reread their papers, noting any comment that had been written on them and to consider how they might improve their next essay. Students without comments were instructed to reread their papers, looking for ways to improve their writing.

After students had reviewed their compositions, they were given the second topic and asked to write on it. The same procedure was followed for the remainder of the writing experiment. Each writing day, students were given their previous compositions and were asked to reread them, considering any comments that had been made.

After the last composition had been written, students were asked to complete a questionnaire to assess their attitudes toward the writing experiment (see Appendix B).
Analysis of the Data

Compositions I and IV were read to assess the mean difference in scores of quantity and quality. Attitude questionnaires were tabulated to determine whether the attitudes of students were generally positive or negative toward the writing experiment.

**Quantity**

A T-unit, according to Hunt, is a minimal terminable unit made up of one main clause plus the subordinate clauses attached to it or embedded within it (2, p. 49). Compositions I and IV were marked for number of T-units, and the mean difference in number of T-units for papers I and IV was computed for each student. Perfect inter-scorer reliability was established for number of T-units.

**Quality**

The quality of papers I and IV was rated individually by three qualified English teachers who had had experience in teaching junior English. Before rating the compositions in the experiment, the raters studied the rating procedure and agreed on its use. Ten compositions from another school were used to establish reliability. Inter-rater correlations of .70, .72, and .82 were established for quality. The total scores of each rater were then averaged to give a single index of quality for each paper. The mean difference between papers I and IV was then computed.
Attitudes

At the end of the experiment students were given a questionnaire to determine their attitudes toward the writing experiment (see Appendix B). The questionnaire was adapted from one used by Taylor (4). Eleven questions which could be answered negatively or positively were developed. For the purpose of this study, any student who responded positively to six or more questions was considered as having generally favorable attitudes toward the experiment. Any student who responded negatively to six or more questions was considered as having negative attitudes toward the experiment.

Statistical Treatment

A one-way analysis of variance was computed to test Hypotheses I, II, III, and IV. The analyses tested the mean differences for combined subjects and separately for high, middle, and low ability level subjects. Because significant $F$ values of .05 or less were found, the Duncan's New Multiple Range Test was used to find where the significant differences existed.

Parts A and B of Hypotheses V, VI, and VII were tested by a $t$ test. Part C of Hypotheses V, VI, and VII and Hypotheses VIII and IX were tested using Chi Square. The .05 level of confidence was required for significance.
CHAPTER BIBLIOGRAPHY


CHAPTER IV

PRESENTATION AND ANALYSIS OF DATA

The basic purpose of this study was to determine the relative effectiveness of three types of written comments on subsequent student compositions. In this chapter, data gathered on quantity, quality, and attitudes will be presented and discussed. Results for total subjects, high-ability subjects, middle-ability subjects, and low-ability subjects will be considered separately.

Findings for Total Subjects

Research Hypotheses I-A, II-A, III-A, and IV-A predicted findings for the combined subjects. It was anticipated that there would be significant differences among the subjects who were praised, criticized, and received no comment on measures of quantity, quality, and attitude. The significant differences found were expected to show that praise effected the most positive gains and that criticism effected more positive gains than no comment. The means and standard deviations for total subjects on each of the quantitative measures are presented in Table III.

These data were subjected to further analysis to determine in which measures there was a significant difference among the means for the three types of written comment.
TABLE III
MEAN GAINS AND STANDARD DEVIATIONS FOR QUANTITATIVE MEASURES FOR TOTAL SUBJECTS

<table>
<thead>
<tr>
<th>Type of Comment</th>
<th>Measure</th>
<th>Mean</th>
<th>S.D.</th>
<th>Mean</th>
<th>S.D.</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Praise</td>
<td>Criticism</td>
<td>No Comment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quantity</td>
<td>Mean</td>
<td>-1.8</td>
<td>7.44</td>
<td>-5.83</td>
<td>8.84</td>
<td>-4.71</td>
<td>7.40</td>
</tr>
<tr>
<td>Quality</td>
<td>Mean</td>
<td>.26</td>
<td>2.42</td>
<td>.20</td>
<td>2.86</td>
<td>-.11</td>
<td>2.74</td>
</tr>
</tbody>
</table>

Simple analysis of variance was used to test for significance.

Hypothesis I-A stated that there would be a significant difference among combined subjects on quantity. Table IV shows the analysis of variance data for number of T-units.

TABLE IV
SUMMARY OF ANALYSIS OF VARIANCE OF NUMBER OF T-UNITS FOR TOTAL SUBJECTS USING THREE TYPES OF CRITICISM

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sums of Squares</th>
<th>df</th>
<th>Variance Estimate</th>
<th>F Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between subjects</td>
<td>404.22</td>
<td>2</td>
<td>202.11</td>
<td>3.14*</td>
</tr>
<tr>
<td>Within subjects</td>
<td>8743.16</td>
<td>136</td>
<td>64.29</td>
<td></td>
</tr>
<tr>
<td>Total subjects</td>
<td>9147.38</td>
<td>138</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant at the .05 level of confidence.

In each of the summaries of analysis of variance, the analysis was made to determine whether the mean difference between the groups tested was larger than the difference among individuals regardless of group assignment. In order for one
treatment group to prove more effective than another treatment group, the mean difference between groups had to be significantly larger than the mean difference among individuals.

The F ratio of 3.14 reveals that the difference among the groups could have occurred by chance only five in 100 times. The Duncan's New Multiple Range Test was used for further analysis of the data.

Table V shows that the number of T-units differed significantly at the .05 level of confidence when comparing the

| TABLE V |
| DUNCAN'S MULTIPLE RANGE TEST OF SIGNIFICANCE BETWEEN PAIRS OF MEANS FOR NUMBER OF T-UNITS FOR TOTAL SUBJECTS |

<table>
<thead>
<tr>
<th>Initial Groups</th>
<th>Ranked Means</th>
<th>Mean Difference</th>
<th>Range Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Praised-Criticized</td>
<td>(-1.81)-(-5.83)</td>
<td>4.02</td>
<td>3.45*</td>
</tr>
<tr>
<td>Praised-No Comment</td>
<td>(-1.81)-(-4.71)</td>
<td>2.90</td>
<td>3.31</td>
</tr>
<tr>
<td>No Comment-Criticized</td>
<td>(-4.71)-(-5.83)</td>
<td>1.12</td>
<td>3.31</td>
</tr>
</tbody>
</table>

*Significant at .05 level of confidence.

praised and criticized groups. There was more difference between the praised and no comment than between the no comment and criticized groups, but neither was statistically significant at the .05 level of confidence.

Hypothesis III-A stated that there would be a significant difference among combined subjects on quality. Table VI shows the analysis of variance data for quality ratings.
TABLE VI
SUMMARY OF ANALYSIS OF VARIANCE OF QUALITY RATINGS FOR TOTAL SUBJECTS USING THREE TYPES OF CRITICISM

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sums of Squares</th>
<th>df</th>
<th>Variance Estimate</th>
<th>F Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between subjects</td>
<td>3.58</td>
<td>2</td>
<td>1.79</td>
<td>.24*</td>
</tr>
<tr>
<td>Within subjects</td>
<td>998.03</td>
<td>136</td>
<td>7.33</td>
<td></td>
</tr>
<tr>
<td>Total subjects</td>
<td>1001.61</td>
<td>138</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Not significant.

Hypothesis I-B can be rejected since the F ratio is not significant. The statistical analysis supports the conclusion that there was no significant difference on quality ratings among subjects who received praise, criticism, and no comment.

Findings for High Ability Subjects

Research Hypotheses I-B, II-B, III-B, and IV-B predicted findings for high-ability subjects. It was anticipated that there would be significant differences among the subjects who were praised, criticized, and received no comment on measures of quantity, quality, and attitude. The significant differences found were expected to show that praise effected the most positive gains and that criticism effected more positive gains than no comment. The means and standard deviations for high-ability-level subjects on each of the quantitative measures are presented in Table VII.
TABLE VII
MEAN DIFFERENCES AND STANDARD DEVIATIONS FOR QUANTITATIVE MEASURES FOR HIGH-ABILITY SUBJECTS

<table>
<thead>
<tr>
<th>Measure</th>
<th>Type of Comment</th>
<th>Praise</th>
<th>Criticism</th>
<th>No Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
<td>Mean</td>
<td>S.D.</td>
</tr>
<tr>
<td>Quantity</td>
<td>- .44</td>
<td>9.29</td>
<td>-8.33</td>
<td>9.04</td>
</tr>
<tr>
<td>Quality</td>
<td>-1.28</td>
<td>2.66</td>
<td>.15</td>
<td>3.41</td>
</tr>
</tbody>
</table>

These data were subjected to further analysis to determine in which measures there was a significant difference among the means for the three types of written comment. Simple analysis of variance was used to test for significance.

Hypothesis I-B stated that there would be a significant difference among high-ability subjects on number of T-units. Table VIII shows the analysis of variance data for number of T-units.

TABLE VIII
SUMMARY OF ANALYSIS OF VARIANCE OF NUMBER OF T-UNITS FOR HIGH-ABILITY SUBJECTS USING THREE TYPES OF CRITICISM

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sums of Squares</th>
<th>df</th>
<th>Variance Estimate</th>
<th>F Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between subjects</td>
<td>482.87</td>
<td>2</td>
<td>241.43</td>
<td>3.41*</td>
</tr>
<tr>
<td>Within subjects</td>
<td>3044.87</td>
<td>43</td>
<td>70.81</td>
<td></td>
</tr>
<tr>
<td>Total subjects</td>
<td>3527.74</td>
<td>45</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant at .05 level of confidence.
The F ratio of 3.41 reveals that the difference among the groups could have occurred by chance only 5 in 100 times. The Duncan's New Multiple Range Test was used for further analysis of the data.

Table IX shows that the number of T-units differed significantly at the .05 level of confidence when comparing the praised and criticized groups. There was more difference between the praised and no comment groups than between the no comment and criticized groups, but neither was statistically significant at the .05 level of confidence.

Hypothesis III-B stated that there would be a significant difference among high-ability subjects on quality. Table X shows the analysis of variance data for quality ratings.

Since the F ratio is not significant, the statistical analysis supports the conclusion that there was no significant difference among subjects who received praise, criticism, and no comment on quality ratings.

**TABLE IX**

DUNCAN'S MULTIPLE RANGE TEST OF SIGNIFICANCE BETWEEN PAIRS OF MEANS FOR NUMBER OF T-UNITS FOR HIGH-ABILITY SUBJECTS

<table>
<thead>
<tr>
<th>Initial Groups</th>
<th>Ranked Means</th>
<th>Mean Difference</th>
<th>Range Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Praised-Criticized</td>
<td>(-.44)-(-8.33)</td>
<td>7.90</td>
<td>6.43*</td>
</tr>
<tr>
<td>Praised-No Comment</td>
<td>(-.44)-(-4.40)</td>
<td>3.96</td>
<td>6.11</td>
</tr>
<tr>
<td>No Comment-Criticized</td>
<td>(-4.40)-(-8.33)</td>
<td>3.93</td>
<td>6.21</td>
</tr>
</tbody>
</table>

*Significant at the .05 level of confidence.
TABLE X

SUMMARY OF ANALYSIS OF VARIANCE OF QUALITY RATINGS FOR HIGH-ABILITY SUBJECTS USING THREE TYPES OF CRITICISM

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sums of Squares</th>
<th>df</th>
<th>Variance Estimate</th>
<th>F Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between subjects</td>
<td>10.88</td>
<td>2</td>
<td>5.43</td>
<td>.59*</td>
</tr>
<tr>
<td>Within subjects</td>
<td>395.86</td>
<td>43</td>
<td>9.20</td>
<td></td>
</tr>
<tr>
<td>Total subjects</td>
<td>406.74</td>
<td>45</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*No significant difference.

Findings for Middle-Ability-Level Subjects

Research Hypotheses I-C, II-C, III-C, and IV-C predicted findings for middle-ability subjects. It was anticipated that there would be significant difference among the subjects who were praised, criticized, and received no comment on measures of quantity, quality, and attitude. The significant differences were expected to show that praise effected the most positive gains and that criticism effected more positive gains than no comment. The mean gains and standard deviations for middle-ability-level subjects on each of the quantitative measures are presented in Table XI.

These data were subjected to further analysis to determine in which measures there was a significant difference among the means for the three types of written comment. Simple analysis of variance was used to test for significance. Table XII shows these data.
TABLE XI
MEAN DIFFERENCES AND STANDARD DEVIATIONS FOR QUANTITATIVE MEASURES FOR MIDDLE-ABILITY SUBJECTS

<table>
<thead>
<tr>
<th>Measure</th>
<th>Type of Comment</th>
<th>Praise</th>
<th>Criticism</th>
<th>No Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
<td>Mean</td>
<td>S.D.</td>
</tr>
<tr>
<td>Quantity</td>
<td>-1.00</td>
<td>6.97</td>
<td>-1.69</td>
<td>7.52</td>
</tr>
<tr>
<td>Quality</td>
<td>-0.28</td>
<td>2.29</td>
<td>0.86</td>
<td>2.62</td>
</tr>
</tbody>
</table>

TABLE XII
SUMMARY OF ANALYSIS OF VARIANCE OF NUMBER OF T-UNITS FOR MIDDLE-ABILITY SUBJECTS USING THREE TYPES OF CRITICISM

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sums of Squares</th>
<th>df</th>
<th>Variance Estimate</th>
<th>F Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between subjects</td>
<td>24.77</td>
<td>2</td>
<td>12.39</td>
<td>.25*</td>
</tr>
<tr>
<td>Within subjects</td>
<td>2111.84</td>
<td>43</td>
<td>49.11</td>
<td></td>
</tr>
<tr>
<td>Total subjects</td>
<td>2136.61</td>
<td>45</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*No significant difference.

The .25 F ratio is not significant at the .05 level of confidence; thus, there was no significant difference among middle-ability-level subjects who were praised, criticized, or received no comment in number of T-units.

Hypothesis III-C stated that there would be a significant difference among middle-ability subjects on quality. Table XIII shows the analysis of variance data for quality ratings.
TABLE XIII

SUMMARY OF ANALYSIS OF VARIANCE OF QUALITY RATINGS FOR MIDDLE-ABILITY SUBJECTS USING THREE TYPES OF CRITICISM

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sums of Squares</th>
<th>df</th>
<th>Variance Estimate</th>
<th>F Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between subjects</td>
<td>10.77</td>
<td>2</td>
<td>5.38</td>
<td></td>
</tr>
<tr>
<td>Within subjects</td>
<td>267.40</td>
<td>43</td>
<td>6.22</td>
<td>.87*</td>
</tr>
<tr>
<td>Total subjects</td>
<td>278.17</td>
<td>45</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*No significant difference.

Since the F ratio is not significant at the .05 level of confidence, the statistical analysis rejects the hypothesis that there would be significant differences among subjects who received praise, criticism, and no comment on quality ratings.

Findings for Low-Ability-Level Subjects

Research Hypotheses I-D, II-D, III-D, and IV-D hypothesized findings for low-ability subjects. It was anticipated that there would be significant differences among the subjects who were praised, criticized, and received no comment on measures of quantity, quality, and attitude. The significant differences were expected to show that praise effected the most positive gains and that criticism effected more positive gains than no comment. The means and standard deviations for low-ability-level subjects on each of the quantitative measures are presented in Table XIV.
TABLE XIV

MEAN DIFFERENCES AND STANDARD DEVIATIONS FOR QUANTITATIVE MEASURES FOR LOW-ABILITY-LEVEL SUBJECTS

<table>
<thead>
<tr>
<th>Measure</th>
<th>Type of Comment</th>
<th>Praise</th>
<th>Criticism</th>
<th>No Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
<td>Mean</td>
<td>S.D.</td>
</tr>
<tr>
<td>Quantity</td>
<td>-3.94</td>
<td>4.88</td>
<td>-7.63</td>
<td>8.37</td>
</tr>
<tr>
<td>Quality</td>
<td>-0.26</td>
<td>1.90</td>
<td>-0.41</td>
<td>2.34</td>
</tr>
</tbody>
</table>

These data were subjected to further analysis to determine in which measures there was a significant difference among the means for the three types of written comment. Simple analysis of variance was used to test for significance.

Hypothesis I-D stated that there would be a significant difference among low-ability-level subjects on quantity. Table XV shows the analysis of variance data for number of T-units.

TABLE XV

SUMMARY OF ANALYSIS OF VARIANCE OF NUMBER OF T-UNITS FOR LOW-ABILITY-LEVEL SUBJECTS USING THREE TYPES OF CRITICISM

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sums of Squares</th>
<th>df</th>
<th>Variance Estimate</th>
<th>F Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between subjects</td>
<td>122.34</td>
<td>2</td>
<td>61.17</td>
<td>.92*</td>
</tr>
<tr>
<td>Within subjects</td>
<td>2923.62</td>
<td>44</td>
<td>66.46</td>
<td></td>
</tr>
<tr>
<td>Total subjects</td>
<td>3045.96</td>
<td>46</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Not significant.
Since the F ratio is not significant at the .05 level of confidence, the statistical analysis rejects the hypothesis that there would be a significant difference among subjects who received praise, criticism, and no comment on number of T-units.

Hypothesis III-D stated that there would be a significant difference among low-ability subjects on quality. Table XVI shows the analysis of variance data for quality ratings.

### TABLE XVI

**SUMMARY OF ANALYSIS OF VARIANCE ON QUALITY RATINGS FOR LOW-ABILITY SUBJECTS USING THREE TYPES OF CRITICISM**

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sums of Squares</th>
<th>df</th>
<th>Variance Estimate</th>
<th>F Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between subjects</td>
<td>12.50</td>
<td>2</td>
<td>6.25</td>
<td></td>
</tr>
<tr>
<td>Within subjects</td>
<td>249.86</td>
<td>44</td>
<td>5.68</td>
<td>1.10*</td>
</tr>
<tr>
<td>Total subjects</td>
<td>262.36</td>
<td>46</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*No significant difference.

Since the F ratio is not significant at the .05 level of confidence, the statistical analysis rejects the hypothesis that there would be significant differences among subjects who received praise, criticism, and no comment on quality ratings.

**Findings for Differences Between Sexes**

Research Hypotheses V-A and B, VI-A and B, and VII-A and B stated that there would be no significant differences
between sexes on measures of quantity and quality. Table XVII shows a summary of the $t$ test between the means of males and females who were praised.

**TABLE XVII**

**SUMMARY OF $t$-TEST FOR PRAISED MALES AND FEMALES ON QUANTIFIED DATA**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Males</th>
<th></th>
<th>Females</th>
<th></th>
<th>Fisher's $t$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
<td>Mean</td>
<td>S.D.</td>
<td></td>
</tr>
<tr>
<td>Quantity</td>
<td>-2.76</td>
<td>5.29</td>
<td>-.73</td>
<td>9.19</td>
<td>-.92*</td>
</tr>
<tr>
<td>Quality</td>
<td>.20</td>
<td>2.35</td>
<td>.31</td>
<td>2.49</td>
<td>-.15*</td>
</tr>
</tbody>
</table>

*No significant difference.*

Since the $t$ value was not significant at the .05 level of confidence, one could not conclude that there was a significant difference between males and females who were praised as prescribed in this study. Table XVIII shows a summary of the $t$ test between the means of males and females who were criticized.

**TABLE XVIII**

**SUMMARY OF $t$-TEST FOR CRITICIZED MALES AND FEMALES ON QUANTIFIED DATA**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Males</th>
<th></th>
<th>Females</th>
<th></th>
<th>Fisher's $t$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
<td>Mean</td>
<td>S.D.</td>
<td></td>
</tr>
<tr>
<td>Quantity</td>
<td>-6.35</td>
<td>9.83</td>
<td>-5.36</td>
<td>8.06</td>
<td>-.44*</td>
</tr>
<tr>
<td>Quality</td>
<td>.32</td>
<td>3.52</td>
<td>.13</td>
<td>2.30</td>
<td>.22*</td>
</tr>
</tbody>
</table>

*No significant difference.*
The $t$ values in the test between criticized males and females indicates that there were no significant differences between sexes on quantity or quality. Table XIX shows a summary of the $t$ test between the means of males and females who received no comment.

**TABLE XIX**

**SUMMARY OF $t$-TEST FOR NO COMMENT MALES AND FEMALES ON QUANTIFIED DATA**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Males</th>
<th>Females</th>
<th>Fisher's $t$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
<td>Mean</td>
</tr>
<tr>
<td>Quantity</td>
<td>-3.22</td>
<td>5.74</td>
<td>-5.70</td>
</tr>
<tr>
<td>Quality</td>
<td>- .22</td>
<td>2.40</td>
<td>- .04</td>
</tr>
</tbody>
</table>

*No significant difference.

Research Hypotheses V-A and B, VI-A and B, and VII-A and B were all supported since no significant difference was found between males and females in any treatment group on measures of quantity or quality.

**Findings for Attitude Differences**

Research Hypotheses V-C, VI-C, VII-C, VIII, and IX predicted differences among subjects on attitudes toward composition. It was anticipated that subjects who were praised would have the most positive attitudes, that subjects who were criticized would have more positive attitudes than those who received no comment. It was further predicted
that there would be no significant difference among females
and males in their attitudes toward composition.

Research Hypothesis V-C predicted that among the group
receiving praise there would be no significant difference
between sexes in their attitudes toward composition. Table
XX shows a summary of the Chi Square test between the males
and females who were praised.

**TABLE XX**

SUMMARY OF CHI SQUARE TEST FOR DIFFERENCE BETWEEN ATTITUDES
OF MALES AND FEMALES WHO WERE Praised

<table>
<thead>
<tr>
<th>Group</th>
<th>Positive Attitudes</th>
<th>Negative Attitudes</th>
<th>Total</th>
<th>Chi Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed</td>
<td>24.00</td>
<td>1.00</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>23.93</td>
<td>1.06</td>
<td></td>
<td>.009*</td>
</tr>
<tr>
<td>Females:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed</td>
<td>21.00</td>
<td>1.00</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>21.06</td>
<td>.93</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Not significant.

Research Hypothesis V-C can be accepted since the Chi
Square of .009 was not significant. There was no difference
in the attitudes of males and females who were praised.

Research Hypothesis VI-C predicted that among the group
receiving criticism there would be no significant difference
in the attitudes of males and females toward composition.
Table XXI shows a summary of the Chi Square test between the
males and females who were criticized.
TABLE XXI

SUMMARY OF CHI SQUARE TEST FOR DIFFERENCE BETWEEN ATTITUDES OF MALES AND FEMALES WHO WERE CRITICIZED

<table>
<thead>
<tr>
<th>Group</th>
<th>Positive Attitudes</th>
<th>Negative Attitudes</th>
<th>Total</th>
<th>Chi Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed</td>
<td>11.00</td>
<td>8.00</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>10.91</td>
<td>8.09</td>
<td></td>
<td>.003*</td>
</tr>
<tr>
<td>Females:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed</td>
<td>16.00</td>
<td>12.00</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>16.09</td>
<td>11.91</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Not significant.

Research Hypothesis VI-C can be accepted since the Chi Square of .003 was not significant. There was no significant difference in the attitudes of males and females who were criticized.

Research Hypothesis VII-C predicted that among the group receiving no comment there would be no significant difference between sexes in their attitudes toward composition. Table XXII shows a summary of the Chi Square test between the males and females who received no comment.

Research Hypothesis VII-C can be accepted since the Chi Square of 3.38 was not significant. There was no significant difference in the attitudes of males and females who received no comment.

Hypothesis VIII predicted that subjects who were praised would show attitudes which were significantly more
**TABLE XXII**

SUMMARY OF CHI SQUARE TEST FOR DIFFERENCE BETWEEN ATTITUDES OF MALES AND FEMALES WHO RECEIVED NO COMMENT

<table>
<thead>
<tr>
<th>Group</th>
<th>Positive Attitudes</th>
<th>Negative Attitudes</th>
<th>Total</th>
<th>Chi Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed</td>
<td>13</td>
<td>5</td>
<td>18</td>
<td>3.38*</td>
</tr>
<tr>
<td>Expected</td>
<td>10</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed</td>
<td>12</td>
<td>15</td>
<td>27</td>
<td>3.38*</td>
</tr>
<tr>
<td>Expected</td>
<td>15</td>
<td>12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Not significant.

positive than those who received criticism. Table XXIII shows a summary of the Chi Square test between those who were praised and those who were criticized.

**TABLE XXIII**

SUMMARY OF CHI SQUARE TEST FOR DIFFERENCES IN ATTITUDES BETWEEN THOSE WHO WERE PRAISED AND THOSE WHO WERE CRITICIZED

<table>
<thead>
<tr>
<th>Group</th>
<th>Positive Attitudes</th>
<th>Negative Attitudes</th>
<th>Total</th>
<th>Chi Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed</td>
<td>42</td>
<td>2</td>
<td>47</td>
<td>19.23*</td>
</tr>
<tr>
<td>Expected</td>
<td>36</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed</td>
<td>27</td>
<td>20</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>36</td>
<td>11</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant at .001 level of confidence.
Hypothesis VIII may be accepted at the .001 level of confidence indicating that those who were praised had significantly more positive attitudes than those who were criticized.

Hypothesis IX predicted that subjects who were criticized would show attitudes which were significantly more positive than those who received no comment. Table XXIV shows a summary of the Chi Square test between those who were criticized and those who received no comment.

### TABLE XXIV

**SUMMARY OF CHI SQUARE TEST FOR DIFFERENCES IN ATTITUDES BETWEEN THOSE WHO WERE CRITICIZED AND THOSE WHO RECEIVED NO COMMENT**

<table>
<thead>
<tr>
<th>Group</th>
<th>Positive Attitudes</th>
<th>Negative Attitudes</th>
<th>Total</th>
<th>Chi Square</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Males:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed</td>
<td>27</td>
<td>20</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>26.57</td>
<td>20.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Females:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed</td>
<td>25</td>
<td>20</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>25.43</td>
<td>19.57</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Not significant.*

Hypothesis IX, that there would be significant differences in the attitudes of those criticized and those who received no comment, must be rejected since the Chi Square of .033 did not reach significance.
Table XXV presents a summary of the Chi Square test between those praised and those who received no comment. A comparison between the Chi Square values of the praised-criticized groups can be made with the criticized-no comment groups.

**TABLE XXV**

**SUMMARY OF CHI SQUARE TEST FOR DIFFERENCES IN ATTITUDES BETWEEN THOSE WHO WERE PRAISED AND THOSE WHO RECEIVED NO COMMENT**

<table>
<thead>
<tr>
<th>Group</th>
<th>Positive Attitude</th>
<th>Negative Attitude</th>
<th>Total</th>
<th>Chi Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed</td>
<td>45</td>
<td>2</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>35.76</td>
<td>11.24</td>
<td></td>
<td>20.41*</td>
</tr>
<tr>
<td>Females:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed</td>
<td>25</td>
<td>20</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>34.24</td>
<td>10.76</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant at the .001 level of confidence.

The Chi-Square of the praised-no comment test is larger than the Chi-Square of the praised-criticized test. The difference indicates that there was even more difference between the praised and no comment groups than between the praised and criticized groups.

**Summary of Findings**

Detailed analysis of the statistical treatments for total subjects, high-ability subjects, middle-ability subjects, and low-ability subjects has been presented in
Chapter IV. Table XXVI draws together the findings to show where significant differences were found.

**TABLE XXVI**

**SUMMARY OF FINDINGS FOR QUANTITATIVE DATA FOR THREE TREATMENT GROUPS**

<table>
<thead>
<tr>
<th>Measures</th>
<th>Total Subjects</th>
<th>High-Ability Subjects</th>
<th>Middle-Ability Subjects</th>
<th>Low-Ability Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant at the .05 level of confidence.

In addition to the analysis of the effect of written comment on quantity and quality, Chapter IV presented statistical data on the effects of written comments on the attitudes of students toward writing. Table XXVII presents a summary of those findings.

**TABLE XXVII**

**SUMMARY OF FINDINGS FOR ATTITUDES OF THREE TREATMENT GROUPS**

<table>
<thead>
<tr>
<th>Groups</th>
<th>Significant Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Praised Males vs Praised Females</td>
<td>N.S.</td>
</tr>
<tr>
<td>Criticized Males vs Criticized Females</td>
<td>N.S.</td>
</tr>
<tr>
<td>No Comment Males vs No Comment Females</td>
<td>N.S.</td>
</tr>
<tr>
<td>Total Praised vs Total Criticized</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Total Praised vs Total No Comment</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Total Criticized vs Total No Comment</td>
<td>N.S.</td>
</tr>
</tbody>
</table>
CHAPTER V

SUMMARY, CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS

This research study was an investigation of the effects of praise, criticism, and no comment on the expository compositions of eleventh-grade students. The effects of teacher comments were measured on quantity, quality, and on the attitudes of students toward writing. This chapter contains a summary of the methods and procedures used to obtain the data, a review of the findings, conclusions, implications, and recommendations for educational practice and future research.

Summary of Methods and Procedures

One hundred thirty-nine students from junior English classes were divided into high, middle, and low ability groups on the results of the California Short-Form Test of Mental Maturity. Once the experiment began students were allotted fifty minutes to write on each of four topics. After each writing trial the experimenter wrote comments on each paper appropriate to the group to which the student was assigned. One third of the students received comments of praise on each paper they wrote. One third of the experimental group received criticism and suggestions for
improvement. The remaining one third received only a check mark that their paper had been read. No grades were given on any paper. Before each successive writing trial, the student's previous composition was returned to him, and he was asked to re-read his paper and consider the comments that had been made. If he had not received comments, he was asked to re-read his paper looking for ways he could improve his next essay. After the last writing trial papers I and IV were analyzed for quantity by computing the number of T-units (1). Mean difference in quality was determined by a panel of three English teachers using an evaluative scale developed by Diederich (see Appendix A). The attitudes of students toward writing were determined by a questionnaire (see Appendix B).

Nine hypotheses had been formulated to predict where significant differences among the groups would appear for total subjects, high-ability subjects, middle-ability subjects, and low-ability subjects. In addition, data were analyzed to test differences between sexes. Analysis of variance and Duncan's New Multiple Range Test were used to establish statistical significance for differences in quantity and quality. Fisher's t test and Chi Square were used to test for significant differences in attitudes.
Summary of the Findings

In Chapter IV where each of the hypotheses was considered separately, tables are given for the analysis of variance computations. If the F ratio indicated that the three treatment groups differed significantly, the data were investigated further by Duncan's New Multiple Range Test to find exactly where the difference existed. Fisher's t test was used to test the difference between sexes within each treatment group. Chi Square was used to test the difference in attitudes. Results of the investigation can be summarized in the following findings:

1. For total subjects the mean difference in number of T-units, as a measure of quantity, differed significantly.

2. For total subjects there was no significant difference among the three treatment groups in quality ratings.

3. For high-ability subjects the mean difference in the number of T-units differed significantly.

4. For high-ability subjects there was no significant difference among the three treatment groups in quality ratings.

5. There was no significant difference in mean difference of quantity or quality for middle- or low-ability-level subjects.

6. There was no significant difference in attitudes of males and females.
7. There was a highly significant difference in attitudes between those who were praised and those who received criticism.

8. There was a highly significant difference in attitude between those who were praised and those who received no comment.

9. There was no significant difference in attitude between those who were criticized and those who received no comment.

Conclusions

Analyses of the findings of this study of the effects of written comment on expository compositions of eleventh-grade students led to the following conclusions:

1. The quantity that a student writes is affected by the type of written comment that a teacher may use. The quantity of high-ability-level subjects is more affected by teacher comments than the quantity of middle- or low-ability-level students.

2. Over a short period of time the quality of a student's writing is not significantly affected by written comments.

3. There are no significant differences between sexes in the effect of written comment on quantity, quality, or attitudes.

4. The results of this study support the learning principle that praise is more effective than criticism or
no comment in developing positive attitudes among students. The lack of significant difference between the attitudes of those who received no comment and those who were criticized might be accounted for by the fact that students whose papers lacked any element of praise perceived their treatment as a form of rejection or punishment. Both the criticized and no comment groups, perhaps, saw their treatment as a form of rejection and felt that the comments were an indictment of unworthiness.

Implications

The following implications were drawn from the findings and conclusions of this study:

1. Negative criticism or no feedback causes students to write less than students who are praised; therefore, teachers should use praise (positive reinforcement) to encourage students to write more.

2. On measures of quantity and attitudes, high-ability-level students were more affected by written comments than were middle- or low-ability-level students. The notion that high-ability-level students will function "in spite of the teacher" was disputed by the findings in this experiment.

3. Teachers will not find that the type of written comment alone, over a short period of time, will make a significant difference in over-all quality of students' writing. Improved quality may be effected by teacher comments over a longer period of time. Improved quality may also be
dependent on other variables such as use of models, group
instruction, or maturation.

4. There is no significant difference in the effects
comments have on male and female students. Whatever differ-
ences in reaction to adult approval or disapproval may have
existed in younger students seems to be minimized in the
high school years. Despite Sweet's findings that ninth-grade
boys were more affected by teacher comments than were girls,
these findings do not warrant such a conclusion (2).

5. Attitudes are significantly affected by teachers'
written comments. In this study praised students had more
positive attitudes than students who were criticized or re-
ceived no comment. If teachers are to build positive atti-
tudes toward writing, comments of praise for successive
approximations of an ideal performance seem to be essential.

Recommendations

1. Since this experiment was conducted over a four-
week period, the effects of written comments on quality of
composition over a longer period of time need further investi-
gation. Would quality be affected by written comments if
the comments were given over a longer period of time?

2. In addition to groups who are given only praise,
given only negative criticism, or given no comment at all,
further research including a group who receive both praise
and criticism needs to be made.
3. Further studies to determine the effect of isolated variables such as the use of models, maturation, and classroom instruction on quality should be made.

4. Studies to compare the effects of teacher comments and peer comments on subsequent student improvement need to be made.

5. In further research dealing with quantity of composition, some measure which would give weight to the number of subordinate clauses should be utilized.

6. Studies to determine the relative effects of a variety of reinforcements should be made.
CHAPTER BIBLIOGRAPHY


2. Sweet, Roger C., Educational Attainment and Attitudes Toward School as a Function of Feedback in the Form of Teachers' Written Comments, Madison, Wisconsin, November, 1966, Educational Resources Information Center (ED 015 163).
APPENDIX A

CRITERIA FOR JUDGING QUALITY OF COMPOSITION

Categories of Comments

I. IDEAS

Relevance
Concise-wordy
Clarity of Ideas
Development
Too Brief or Long
Persuasiveness
Ending
Generality . . . . . . . . . . . . . . . 1-5 points

II. FORM

Spelling
Clarity of Expression
Organization
Coherence of Ideas
Reader Agreement
Analysis
Maturity . . . . . . . . . . . . . . . . 1-5 points

III. FLAVOR

Quality of Ideas
Style (general)
Mechanics (general)
Originality
Interest
Beginning
Sincerity
Intro. & Illus. . . . . . . . . . . . . . . . . 1-5 points

IV. MECHANICS

Punctuation
Grammar
Sentence Structure
Phrasing, Idiom . . . . . . . . . . . . . . . . . . . . 1-5 points
V. WORDING

General
Word Choices
Logic
Cliches
Jargon-Slang . . . . . . . . . . . . 1-5 points

Total points 5-25
APPENDIX B

QUESTIONNAIRE

Please respond yes or no to each of the following statements by placing a check in the appropriate blank.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I enjoyed writing the essays.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. At first I didn't like writing, but I changed my mind and learned to enjoy writing.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Comments on my paper, or the lack of them, made me feel angry with the grader.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Comments on my paper, or the lack of them, made me feel like trying harder.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Comments on my paper, or the lack of them, made me feel good about my work.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Comments on my paper, or the lack of them, made me feel angry with myself.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. The comments overlooked the best points of my writing.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. I would like to continue writing for this experiment.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. I feel that my writing improved because of the comments that were written on my paper.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. The comments on my paper, or the lack of them, are the kinds I would like all teachers to make on work that I hand in.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. In general, I liked the way my essays were marked.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX C

ESSAY TOPICS

I. Most teen-agers have at least one problem that gives them considerable concern. With some, dating is a problem. With others, getting money or a job is a problem. Some teen-agers worry about their appearance or being popular. Still others worry about their future career or a major problem facing the nation or the world.

What is your main concern at this time? Explain why your concern became important to you and why you consider it worth overcoming. You might also explain how various people or situations contribute to the problem.

II. Newspapers, television, and even popular records keep saying that there is a "generation gap" in the world today. What they mean, of course, is that the younger people and the older people cannot or will not understand one another.

Do you think there is a lack of understanding between younger and older people? Cite some examples to back up your opinion and explain why you think the gap exists.
III. People today are working fewer hours, and some experts predict that by the year 2000, most people will work no more than twenty hours a week. With more free time people must find new and better ways to spend their time.

What are your suggestions? Explain several ways you think people could spend their time profitably and enjoyably.

IV. Happiness has been defined by many great writers. Perhaps you could not accept any of their definitions because what makes one person happy doesn't necessarily make someone else happy.

What do you feel is necessary to make you a happy person? Explain why each point is important.
APPENDIX D

INSTRUCTIONS TO STUDENTS

You have been selected to participate in an important experiment in writing. English teachers are interested in improving how they teach writing, and this experiment may help to improve the teaching of composition and may help to enhance your enjoyment of writing. The results, of course, will depend on your willingness to stick with the experiment and on your willingness to do your best.

The four compositions that you will write in the experiment will be considered a part of your course requirements, and your score on the compositions will be a part of your grade for the six weeks. No grade, however, will appear on the composition until all four have been written and the experiment is concluded. Please do not ask what grade you made on any composition until that time.

Each week you will be given a topic about which you will probably have a great deal of knowledge and one about which you are interested. During the period each Wednesday, you will be asked to write without the aid of help. Before each successive writing period, your papers will be returned to you.

Please do not ask your teacher to comment or advise you regarding your writing. The experiment requires that no
information about the experiment be given until after it is over. At the same time, do not feel that you have to be a good writer to participate. The experiment is more for getting your reaction to the topics than to judge your greatness as a writer.

After the experiment is over, I will come to explain what the experiment was intended to do and to answer any questions you may have. I hope you will cooperate and will find the experiment enjoyable and worthwhile.

At this point students will be asked if they have any questions.
BIBLIOGRAPHY

Books


Hunt, Kellogg W., Grammatical Structures Written at Three Grade Levels, Champaign, Illinois, National Council of Teachers of English, 1965.


Shugrue, Michael F., English in a Decade of Change, New York, Pegasus, 1968.


Articles


Publications of Educational Resources Information Center


Sweet, Robert C., *Educational Attainment and Attitudes Toward School as a Function of Feedback in the Form of Teachers' Written Comments*, University of Wisconsin, Madison, Wisconsin, 1966 (ED 015 163).