DOCUMENT RESUME

ED 032 871 JC 690 345

By-Kivits, Virginia M.
An Experiment in Teaching English Composition Using an Oral Laboratory Approach.
Minnesota Univ., Minneapolis. General College.
Pub Date 69
Note-13p.
Journal Cit-The General College Studies; v5 n6 1968-1969
EDRS Price MF-$0.25 HC-$0.75
Descriptors-*College Freshmen, *Composition Skills (Literary), Grading, *Grammar, *Junior Colleges, *Language Laboratories
Identifiers-Minnesota

Although more college students are enrolled in freshman composition classes than in any other, there is little clarity of purpose in these writing courses. In such a field, where outcomes must be expressed in terms of student writing skills acquired (or not acquired), few well-defined experiments have been reported. This investigator, in an attempt to discriminate among levels of performance in student themes, faced the problem of establishing a reliable evaluation technique. The techniques she finnaly adopted are described in this report. Having set up a grading system, she then sought to learn if one way of teaching grammar was more effective than another. Of two sections in the grammar course, one became the control group, taught in the traditional way, while the other, the experimental group, received three hours of language laboratory instruction and one hour in a recitation-discussion class. Tables show the correlation of rank and raters in grading the themes for each group, and pre- and post-test scores for each class. It appeared that gains were greater in style than in mechanics. The comparisons showed that theme improvement was not appreciable related to the teaching method or to the quarter of instruction. The students did improve their techniques after instruction in grammar and style. The method of grouping showed a significant upward shift in style, but an insignificant one in mechanics. (HH)

UNIVERSITY OF MINNESOTA

Volume V, Number 6 1968-1969

AN EXPERIMENT IN TEACHING ENGLISH COMPOSITION USING AN ORAL LABORATORY APPROACH

by

Virginia M. Kivits

At any given moment during the academic year more students in colleges and universities are probably enrolled in freshman composition courses than in any other subject-matter sequence. And it is also likely that among these myriad students and their mentors there is a higher degree of anguish, despair and frustration than in any other field. Much of the distress felt by both students and teachers of composition arises from a deplorable lack of clarity of purpose in writing courses. It is an educational paradox that in a field that enrolls so many students there is so little agreement about the means of judging the outcomes of instruction. Teachers of composition, like other educators, know that attempts to evaluate teaching effectiveness must be orderly: they know and appreciate the necessity for control groups, for unbiased data gathering, and for establishing objective evaluative criteria. But in a field such as composition, in which outcomes must necessarily be expressed in terms of student writing skills acquired -- or not acquired -- well-designed experiments have rarely been reported. When they have been attempted, studies of the effectiveness of teaching composition have foundered on the shoals of inadequate answers to such questions as "What factors should go into theme grading?" or "Why is this theme better than that one?" This is the type of question considered in the study reported in this issue of The General College Studies.

In an attempt to discover whether one means of teaching grammar is more effective than another, the conductor of the experiment reported here found it necessary to evaluate student writing skills as represented by themes students wrote. In order to discriminate among levels of performance in student themes, the experimenter faced the problem of establishing the reliability of a theme evaluation technique. How such reliability was established -- how, that is, the effort was made to objectify theme grading -- should be of practical as well as theoretical interest to all teachers of composition. Aside from the evaluative techniques described here, this report should be of particular interest to those college writing and grammar teachers who, like the author, feel the need to try and test new techniques in their field.

The author of this report is an Associate Professor in the General College Division of Literature, Writing, and Speech. This report is an abridgement of the author's full account of her experiment, written for her colleagues in the Division.

Editors. G. GORDON KINGSLEY, Coordinator, Student Personnel Services • DAVID L. GIESE, Coordinator of Research • NORMAN W. MOEN, Assistant D

I. Introduction

This report is a summary of an experimental project in which an attempt was made to teach a General College grammar review course using the facilities of a language laboratory. The nature of the content of the grammar course -- one that had for many years been a part of the offerings of General College's Division of Literature, Writing and Speech -- can be gathered from the course description in the 1967-69 General College bulletin:

> GC 30B Fundamentals of Usage and Style
>
> The work in this course includes intensive drill in grammar, spelling, and punctuation. Attention is also given to matters of style, such as subordination, parallel construction, tense consistency, and the placement of modifiers.

As in most institutions that enroll freshmen, the Division's aim in offering such a course is to prepare students for their later enrollment in the freshman composition sequence by providing them with opportunity to refresh their knowledge of informal English usage and to review the fundamentals of written English in sentences.

During the years that Fundamentals of Usage and Style has been a part of the General College's curriculum, several attempts were made to improve the effectiveness of the course as a preparation for college writing. Instructors of 30B through the years noted that, though their students often grew skillful at doing grammar and usage exercises, little of this skill was evident in the writing they did in composition courses. It was in the hope of developing some techniques more relevant to writing that an oral approach to teaching grammar was initiated experimentally in the winter and spring of 1968.

The so-called "oral approach" to teaching language is based on the assumption that all communication skills are related and that proficiency in

one area can be used as a medium of improving another skill. College composition teachers know that though many students are poor writers they may, at the same time, be excellent oral communicators. Teaching writing by means of an oral approach is an attempt to capitalize on the student's strongest verbal skill in a way that will encourage him to become a more accurate and effective writer.

The term "oral approach" as used in this report refers to instructional techniques that make use of language laboratory facilities similar to those now widely used for teaching foreign languages to non-native speakers. The instructional devices provided by a modern language laboratory allow for maximum flexibility of instruction. In separate learning booths, each student makes private use of the prepared material for listening and imitating. Though the techniques used in this experiment were similar to those used in teaching English to foreign students, the materials of instruction had to be developed especially for the needs of the native speaker attempting to integrate his spoken and written language skills. Thus it was necessary, as a preliminary to the project described here, for the author to prepare a special manual and exercise book, the only text used in the experimental sections of the course.

II. Procedures

In order to prepare for evaluation of the techniques employed in the language laboratory approach to teaching grammar, two sections of GC 30B were offered each quarter (Winter and Spring, 1968). One section was designated the "experimental section" and one the "control group." The control group was taught in the more or less traditional manner, using the following texts:

Margaret C. Walters. A Basic Guide to Clear and Correct Writing. Form A. Paperback. Chicago: Scott, Foresman, and Co., 1958.

M. W. Sullivan. Programmed English: A Modern Grammar. Paperback. New York: The Macmillan Co., 1963.

Walters is a traditional grammar text, with explanations, examples, and fill-in exercises. Sullivan is programmed, with weekly or unit tests provided by the author and with a final examination which includes the last unit and a review of the whole text. It does not cover punctuation or composition.

In the control group, students worked fill-in exercises in Walters after the explanations had been reinforced by the teacher, who supplied further examples and explanations. Usually, the exercises were corrected and returned for class discussion, although occasionally the students exchanged and checked each other's papers. In class, the instructor used the blackboard for most of the examples and guided the usual recitation-discussion routine. Every week, students took the unit test provided by the author of the programmed text. At intervals, they took teacher-composed check-tests based on what they were learning from both texts. Upon occasion, they were asked to write their own sentences. At the end of the quarter, they took the final exam composed by the programmed text's author, though those results were not part of the experiment.

The experimental group used the following text:

Virginia M. Kivits. Oral Laboratory Approach to Written English. Mimeographed. University of Minnesota, The General College, 1967.

In the experimental class, time each week was divided between three hours of language laboratory instruction (in which students listened to taped directions while they followed sample sentences and examples in the text, recorded and listened to themselves, and wrote exercises) and a one-hour recitation-discussion class (in which students examined returned and

corrected exercises, asked questions, and were given supplementary explanations).

The experimental class had very little testing during the quarter, the emphasis being primarily upon listening and repeating before writing. Because the exercises were of the sentence and paragraph type, they took more time than fill-in exercises, both in and out of class.

III. Evaluation

In evaluating this project, the outcomes of the oral, laboratory-oriented technique were measured against the results obtained in the traditional approach to teaching grammar. The evaluation consisted of judgments -- by two disinterested but experienced college composition teachers -- of the effectiveness of student performance on written assignments. As a means of gathering data for analysis and evaluation, the instructor asked students in both the experimental and control groups to write impromptu themes at the beginning and at the end of each of the quarters during which the experiment was conducted. These themes were used, first, as the basis of arriving at theme grading reliability and, second, as a measure of student achievement in both the control and experimental sections of the course. The two composition teachers who acted as theme graders met with the instructor to agree upon certain criteria for theme-grading; they also read several groups of themes in order to establish the reliability of the evaluation method. In the discussion below, the data presented in Tables 1 through 4 relates to the establishment of reliability; the data in Tables 5 through 8 relates to the actual judging of student performance.

The theme graders, in consultation with the instructor of the course, agreed to judge student themes on the following bases:

1) Use a numeric scale ranging from 1 (lowest) to 11 (highest).

2) Attempt to use the full range (1-11) of the numeric scale in a "normal" distribution for each group of themes.

3) Assign two grades -- one for Mechanics and one for Style and Effectiveness -- taking into consideration the following criteria:

Mechanics

plural-possessive distinction
general grammar
punctuation
verb tense
splice
pronoun case
agreement
fragment
shift

Style and Effectiveness

Sentence variety and flexibility

position of sentence elements
adjectives
adverbs
phrases
clauses
transitions and other joiners
direct-indirect quotations
interrupting expressions

Emphasis
Active-Passive voice
Sentence rhythm
Verbals

4) Judge style and effectiveness at the sentence level only. (This means that judges were not to consider such matters as theme form, e.g., lack of title, poor format, lack of paragraphing).

5) Omit spelling and capitalization from consideration of Mechanics.

6) Do not take theme length into consideration as a factor in grading effectiveness.

Aside from the rules established above, the theme readers also agreed not to make any marks -- such as correction symbols -- on any student papers and to keep the grades they assigned on separate tabulation sheets.

In the preliminary effort to establish reliability of the evaluation process, the two readers read and assigned numeric grades to two practice sets of themes. The themes were provided by the instructor, who collected them at random from the work of students in both the control and experimental

sections. The themes were distributed to the readers, unidentified in any way, in two folders of seven themes each. In this initial attempt to develop similar evaluation techniques, the graders assigned to each theme a number from one to seven; no ties were allowed in grading each set of themes. Table 1 shows the correlation coefficients[1] between the ranks assigned each paper on both mechanics and style-effectiveness.

TABLE 1

Correlation Between Ranks Assigned in Practice Rounds

Trial	Mechanics	Style
1	.86	.93
2	.36	.04
2*	.70	.70

* adjusted for short-theme consideration

According to the data in this table, the first trial brought high agreement, but the second resulted in disagreement. A consultation about this discrepancy revealed that the cause was mainly the readers' difficulty in ranking the short themes. Therefore, the two readers conferred again in order to try to overcome this problem.

In order to keep checking on the amount of agreement, tabulators computed correlation coefficients immediately after the readers completed the rating of each group of papers. These coefficients indicate the agreement between raters for both factors, mechanics and style, and also between factors for each rater. Table 2 shows both sets of coefficients:

[1]The correlation coefficient is a measure of the amount of agreement between the numbers assigned each paper by each rater, varying between +1 (indicating a perfect agreement) and -1 (indicating identical agreement but in opposite directions). A lack of agreement, either positive or negative, results in a coefficient near zero.

TABLE 2

Correlation between Raters for Each Factor
and between Factors for Each Rater

Group	Number in Group	Between Raters		Between Factors	
		Mechanics	Style	Rater 1	Rater 2
1	23	.79	.75	.91	.89
2	24	.87	.91	.85	.69
3	23	.89	.76	.85	.85
4	23	.69	.59	.89	.77
5	23	.70	.65	.92	.79
6	25	.75	.66	.80	.71
7	23	.81	.69	.91	.65
8	23	.69	.53	.88	.60

Coefficients between raters ranged from a low of .53 to a high of .91, with a median above .70; those between factors were generally higher, ranging from .60 to .92, with a median coefficient near .85. For one rater, mechanics and style appear to be nearly the same (correlations of .80 and above), while for the other, there were more differences. Overall, however, the raters did not differentiate much between mechanics and style. When each group of themes was completed, mean scores were also computed and compared to discover the raters' consistency in using the same scale and to learn whether the groups were similar to, or possibly different from, each other. Table 3 gives these mean scores:

TABLE 3

Mean Scores Assigned by Each Rater for Each Group on the Two Factors

Group	Number in Group	Mechanics Rater 1	Mechanics Rater 2	Style Rater 1	Style Rater 2
1	23	5.7	6.1	6.1	6.1
2	24	6.3	6.3	6.3	6.1
3	23	5.6	6.0	5.5	5.9
4	23	6.2	5.7	6.1	5.6
5	23	7.0	6.3	6.8	6.2
6	25	6.4	6.0	6.2	5.9
7	23	6.2	6.0	6.0	6.0
8	23	6.4	6.2	6.5	6.0
Total	187	6.2	6.1	6.2	6.0

Except for Group 3, which was rated slightly below average on mechanics by both readers, and Group 5, which Rater 1 ranked somewhat above average on mechanics, the group mean scores were in the low sixes, with considerable agreement between the raters. Thus, it seems the grouping had put together themes by writers of about the same ability in writing.

Still another investigation was made to check the reliability and validity of the ranking. This time, the ranks assigned in the trial sessions for the two sets of seven themes were added; the numerics given the same papers when they were graded with the regular groups were also added. Table 4 indicates correlation coefficients between the sum of the ranks in the trials and the numerics in the regular groups:

TABLE 4

Correlation Coefficients between Initial Ranking on Seven-Point Scale and Final Ranking on Eleven-Point Scale

	Mechanics	Style
Trial 1	.90	.92
Trial 2	.74	.90

Again, considerable agreement occurred between the initial ranking when the papers were in the small practice groups and when they were interspersed with the larger groups. When one considers the total findings, he is impressed with the agreement between the raters in the grades they assigned the papers.

Because students were not randomly assigned to the classes, a test was made of the writing ability of students in the four classes. The results are tabulated in Tables 5 and 6. Table 5 gives the initial theme mean scores for each class on both factors, mechanics and style:

TABLE 5

Pre-test Mean Scores for Each Class on Each Factor

Winter Quarter	Number	Mechanics	Style
Experimental	28	5.57	5.36
Control	20	5.45	5.53
Spring Quarter			
Experimental	24	6.60	5.92
Control	21	5.29	5.43

According to this table, the class average in both factors is very close to the middle of the range.

TABLE 6

Analysis of Variance of Pre-test Results for Both Mechanics and Style

Source of Variation	Degrees of Freedom	Mechanics		Style	
		Mean Square	F-Ratio	Mean Square	F-Ratio
Experimental vs. Control	1	45.81	2.03	2.15	.12
Winter vs. Spring	1	23.49	1.04	6.73	.36
Interaction (inconsistency between quarters)	1	32.75	1.45	9.84	.53
Among Individuals	89	22.52		18.55	

Table 6 compares the combined winter-and-spring-quarter experimental groups with the combined winter-and-spring control groups. F-ratios, which are a measure of the difference between groups compared to the average difference within groups, indicate that the four groups (classes) were not significantly different in their ability to write, as judged by the two raters.

Improvement:

Of course, the most important question to be answered is this: Was there improvement in ability to write? Naturally, when the groups of themes were rated, it was hoped that lower ratings would be assigned to the papers written as initial themes than to those written as final ones. Because the initial and final themes for the same writer were always placed in different groups, the difference between initial and final rankings cannot be strictly called a gain score; but it should approximate the student's change in his writing ability.

Scores showing the difference in writing ability from first to final theme in the four class groups are given in Table 7, which also tabulates the t-ratio (a special case of the F-ratio, when only two groups are involved), comparing the mean difference with the average (student) difference within that group:

TABLE 7

Difference Scores and t-Ratios for Each Class on Both Factors

		Mechanics		Style	
Winter Quarter	Number in Group	Difference	t-ratio	Difference	t-ratio
Experimental	28	.89	1.94	1.16	3.90**
Control	20	.75	1.30	.97	2.98**
Spring Quarter					
Experimental	24	.46	.88	.96	3.08**
Control	21	1.04	2.43*	1.09	3.16**

* significant at the .05 level
** significant at the .01 level

Here, it appears that gains are consistently greater in style than in mechanics. All four style t-ratios indicate significantly higher style ratings assigned to final papers than to initial ones, while only one of the mechanics t-ratios indicates a change large enough to be called significant. Nevertheless, all four mechanics t-ratios do indicate a change in a positive direction.

The results of comparing the changes in theme numerics are shown in Table 8 (a two-way disproportionate frequency analysis of variance):

TABLE 8

Analysis of Variance of Difference Results for Mechanics and Style

Source of Variation	Degrees of Freedom	Mechanics		Style	
		Mean Square	F-Ratio	Mean Square	F-Ratio
Experimental vs. Control	1	4.26	.20	.07	.01
Winter vs. Spring	1	1.14	.05	.33	.03
Interaction	1	12.25	.58	2.38	.25
Among Individuals	89	21.21		9.51	

All the F-ratios are sufficiently small to indicate little difference among the groups under study; thus, whether students were in the experimental or control groups during the two quarters under investigation did not matter significantly.

Summary:

It appears from the various comparisons that theme improvement is not appreciably related to either the teaching method or the quarter in which the instruction occurred. However, it is noticeable that students in general changed in writing technique after exposure to instruction in grammar and style. Through the method of grouping, it was found that a significant shift upward

occurred in style of writing; and that while there was also an upward trend in mechanics of English, the shift was not great enough to be regarded as significant.