8-1971

The Effects of Multiple Educational Methodologies on Student Perception of Teacher Performance

James E. Bultman
Western Michigan University

Follow this and additional works at: http://scholarworks.wmich.edu/dissertations

Part of the Teacher Education and Professional Development Commons

Recommended Citation
Bultman, James E., "The Effects of Multiple Educational Methodologies on Student Perception of Teacher Performance" (1971). Dissertations, 3024.
http://scholarworks.wmich.edu/dissertations/3024

This Dissertation—Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
THE EFFECTS OF MULTIPLE EDUCATIONAL METHODOLOGIES ON STUDENT PERCEPTION OF TEACHER PERFORMANCE

by

James E. Bultman

A Dissertation
Submitted to the
Faculty of the Graduate College
in partial fulfillment
of the
Degree of Doctor of Education

Western Michigan University
Kalamazoo, Michigan
August 1971
ACKNOWLEDGEMENTS

Several people have made important and substantial contributions toward the completion of this dissertation. My sincere appreciation is extended to Doctors James Davenport, Charles Helgesen, and James Bosco for their assistance and encouragement while supervising this project. Their unique efforts, as well as those of many others at Western Michigan University, have been essential to this undertaking.

Deserving of a special tribute are the administrators, teachers, and students in the three cooperating school districts. The implementation of this experiment was dependent on their cooperation.

Indebtedness to the Mott Foundation is also recognized. The financial assistance of a generous Fellowship and the opportunity for a stimulating professional experience were instrumental in prompting the pursuit of doctoral study.

Finally, to my wife, Martie, and our children, Matt and Heather, my deepest gratitude for their understanding and patience during this endeavor.

James E. Bultman
BULTMAN, James E., 1941-
THE EFFECTS OF MULTIPLE EDUCATIONAL
METHODOLOGIES ON STUDENT PERCEPTION
OF TEACHER PERFORMANCE.

Western Michigan University, Ed.D., 1971
Education, teacher training

University Microfilms, A XEROX Company, Ann Arbor, Michigan

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ................................................................. ii
LIST OF TABLES ........................................................................ vi
LIST OF FIGURES ..................................................................... viii

CHAPTER

I  INTRODUCTION ................................................................. 1
   Statement of the Problem ................................................. 1
   Objectives of the Study .................................................. 4
   Major Questions Which Were Investigated .................. 8
   Definition of Terms ....................................................... 10
   Importance of the Study ............................................... 11
   Limitations of the Study ............................................... 14
   Assumptions .............................................................. 14
   Organization of the Report .......................................... 15
   Summary ...................................................................... 16

II REVIEW OF RELATED LITERATURE AND RESEARCH ......... 17
   Pedagogical Variety in Teaching .................................. 17
   Viability of Student Rating of Teacher Performance .... 25
   Student Opinion Feedback for Teachers ...................... 38
   Motivation of Log Activity ........................................ 41
   Function of the Consultant .......................................... 45

III RESEARCH DESIGN, SETTING, AND PROCEDURES ........ 50
   Review of the Problem .............................................. 50

iii

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
<table>
<thead>
<tr>
<th></th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>Tabular Summation</td>
<td>141</td>
</tr>
<tr>
<td>D</td>
<td>Feedback</td>
<td>142</td>
</tr>
<tr>
<td>E</td>
<td>Checklist Log</td>
<td>143</td>
</tr>
<tr>
<td>F</td>
<td>Cover Letter</td>
<td>146</td>
</tr>
<tr>
<td>G</td>
<td>Computer Program</td>
<td>147</td>
</tr>
<tr>
<td>H</td>
<td>Student Perception Averages for All Questionnaire Items, Pre and Post</td>
<td>149</td>
</tr>
<tr>
<td>I</td>
<td>Feedback Group Open-Ended Response</td>
<td>153</td>
</tr>
<tr>
<td>J</td>
<td>Log Group Open-Ended Response</td>
<td>154</td>
</tr>
<tr>
<td>K</td>
<td>Planning Group Open-Ended Response</td>
<td>155</td>
</tr>
<tr>
<td>L</td>
<td>Combination Group Open-Ended Response</td>
<td>156</td>
</tr>
<tr>
<td>M</td>
<td>Teacher Self-Perception Averages for All Questionnaire Items, Pre and Post</td>
<td>157</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Overview of the Five Treatment Groups</td>
</tr>
<tr>
<td>2</td>
<td>Reliability of Items on Bryan's Student Opinion Questionnaire</td>
</tr>
<tr>
<td>3</td>
<td>Summary Data for Overall Pretest Measure</td>
</tr>
<tr>
<td>4</td>
<td>Analysis of Variance for Overall Pretest Measure</td>
</tr>
<tr>
<td>5</td>
<td>t-tests of Change Scores (Pre- to Posttest) for Overall Means</td>
</tr>
<tr>
<td>6</td>
<td>t-tests of Change Scores (Pre- to Posttest) for Overall Means, Based on Differences in the Control Group</td>
</tr>
<tr>
<td>7</td>
<td>Summary Data for Overall Posttest Measure</td>
</tr>
<tr>
<td>8</td>
<td>Analysis of Variance for Overall Posttest Measure</td>
</tr>
<tr>
<td>9</td>
<td>Overall Posttest Between Group t-values</td>
</tr>
<tr>
<td>10</td>
<td>Analysis of Variance for Overall Change Score Measure</td>
</tr>
<tr>
<td>11</td>
<td>Summary Data for SOQ Item Nine, &quot;Variety in Teaching,&quot; of the Pretest Measure</td>
</tr>
<tr>
<td>12</td>
<td>Analysis of Variance for SOQ Item Nine, &quot;Variety in Teaching,&quot; of the Pretest Measure</td>
</tr>
<tr>
<td>13</td>
<td>Pretest Between Group t-values for SOQ Item Nine, &quot;Variety in Teaching&quot;</td>
</tr>
<tr>
<td>14</td>
<td>t-tests of Change Scores (Pre- to Posttest) for SOQ Item Nine, &quot;Variety in Teaching&quot;</td>
</tr>
<tr>
<td>15</td>
<td>t-tests of Change Scores (Pre- to Posttest) for SOQ Item Nine, &quot;Variety in Teaching,&quot; Based on Differences in the Control Group</td>
</tr>
<tr>
<td>16</td>
<td>Summary Data for SOQ Item Nine, &quot;Variety in Teaching,&quot; of the Posttest Measure</td>
</tr>
<tr>
<td>TABLE</td>
<td>PAGE</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>17</td>
<td>Analysis of Variance for SOQ Item Nine, &quot;Variety in Teaching,&quot; of the Posttest Measure</td>
</tr>
<tr>
<td>18</td>
<td>Posttest Between Group t-values for SOQ Item Nine, &quot;Variety in Teaching&quot;</td>
</tr>
<tr>
<td>19</td>
<td>Analysis of Variance for SOQ Item Nine, &quot;Variety in Teaching,&quot; of the Change Score Measure</td>
</tr>
<tr>
<td>20</td>
<td>Change Score Between Group t-values for SOQ Item Nine, &quot;Variety in Teaching&quot;</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>18</td>
</tr>
<tr>
<td>2</td>
<td>58</td>
</tr>
<tr>
<td>3</td>
<td>112</td>
</tr>
<tr>
<td>4</td>
<td>113</td>
</tr>
<tr>
<td>5</td>
<td>113</td>
</tr>
</tbody>
</table>

Teacher Role

Research Paradigm

Group B—Incidence of Educational Methodology Employment vs. Time

Group C—Incidence of Educational Methodology Employment vs. Time

Group D—Incidence of Educational Methodology Employment vs. Time
CHAPTER I

INTRODUCTION

Statement of the Problem

Educators have long been concerned about the quality of teacher performance. Building principals in particular generally regard it as one of their primary responsibilities. The 1960 ASCD Yearbook\(^1\) entitled Leadership for Improving Instruction stated the following:

"Although the major focus of all administrative effort is based upon the improvement of instruction, there must be some means of fixing responsibility for the instructional leadership in each building staff. This is the principal's major function."

Since the core of instructional effort takes place at the classroom level it is appropriate that attempts for improvement hinge on the performance of classroom teachers.\(^2\)

More recently the public sector has become concerned about the quality of teacher performance. In Michigan this concern can be traced, in part at least, to the inception in 1965 of collective bargaining which resulted in the spiraling of teachers' salaries.\(^3\)

\(^1\)Hasse, Glen (Chairman), Leadership for Improving Instruction. 1960 Yearbook of the Association for Supervision and Curriculum Development. Washington: The Association, 1960, p. 33.

\(^2\)loc. cit., pp. 97, 164.

Prior to that time it is probable that, although the public did not appreciate poor teaching, they recognized that perhaps the performance was commensurate with the salary. Today, however, with increased taxes necessitated by higher salaries the public has begun to raise some questions. There is considerable sentiment among taxpayers that improved teacher performance should accompany increasing teacher pay raises.\footnote{Cote, William E., "Commission May Jolt Working World of State Teachers." The Grand Rapids Press, LXXVIII (June 7, 1970), 10-E.} \footnote{Cote, William E., "Millikan Plan Would Oust Poor Teachers." The Flint Journal, XCV (December 21, 1969), 1.} The general public knows enough about the systems approach to realize that increased inputs should produce better outputs. Accountability has become the favorite word to describe this relationship between inputs and outputs, resources and results. In a recent publication of the Phi Delta Kappan, devoted specifically to the topic of accountability, Lieberman\footnote{Lieberman, Myron, "An Overview of Accountability." Phi Delta Kappan, LII (December 1970), 194-5.} cited the growing national interest in accountability and its apparent linkage with increased education budgets. It should not be assumed that teacher performance is the only facet of the accountability issue in education; the performances of both policy making personnel and supervisory personnel are also involved. However, Lopez\footnote{Lopez, Felix M., "Accountability in Education." Phi Delta Kappan, LII (December 1970), 232-5.} asserted that teacher performance is perhaps the most basic and the most
pertinent of all accountability phases. Regardless of the relative importance afforded the various phases of accountability in education it is unmistakably clear that the increased emphasis has permeated the area of teacher performance.

Two fundamental possibilities exist for rectifying this teacher performance dilemma. First would be the attempt to improve the performance of present faculties. Secondly, by a process of dismissal and selective recruitment present faculties or parts thereof could be replaced. Admittedly, the latter possibility is a long-range proposition.

With the mandatory enactment in 1964 of the permissive Michigan Tenure Act of 1937, came protection to teachers against unfair dismissal practices. For school districts the effect has been the virtual impossibility of terminating teacher contracts. In describing the situation, state educational columnist Cote concluded:

"State law does give much protection to teachers under the tenure act and various court rulings. The idea of the law was to protect teachers from unreasonable political or other influence. The practical result, though, has been that it is almost impossible now to fire a teacher unless he has been convicted of a crime or caught in a public scandal."

Similarly, the 1960 ASCD Yearbook reported:

"The school administrator will need to be constantly vigilant against the possibility of staffing the system with unfit and incompetent people whom he cannot dismiss because of tenure policy provisions."

Oftentimes lengthy (3 years) and costly ($33,000) legal procedures

---

1Cote, "Millikan Plan Would Oust Poor Teachers," op. cit.

2Hass, op. cit., p. 120.
such as the one in Royal Oak, Michigan\textsuperscript{1} are the results of efforts to dismiss staff members. There has been little help from the teaching profession in discharging inferior teachers. Rather, a general failure of teachers to police their own ranks has been a negative factor in resolving the dilemma.\textsuperscript{2}

Obviously this means that in order to improve the overall competency of staff, school systems will most likely find more viable the possibility of improving their present faculties than dismissing the "undesirables" and recruiting new ones.

It was in an attempt to provide empirical evidence for the possible improvement of teachers in service, as perceived by students, that the present investigation was conducted.

Objectives of the Study

The primary purpose of this study was to determine if teachers could modify student perceptions of their teaching performances by employing multiple educational methodologies in classroom presentations. For example, would it be possible for a teacher, by employing a variety of teaching methods and instructional devices already at his disposal, to modify student perceptions of his teaching performance? A secondary objective was to ascertain the relative

\textsuperscript{1} "State Unit Okays Teacher's Firing." \textit{The Flint Journal}, XCV (November 5, 1969), 53.

\textsuperscript{2} Hummel, Charlton G., "A Right Only if Deserved." \textit{NEA Journal}, XLI (January 1960), 67.
effectiveness of four experimental treatments in helping a teacher to modify student perceptions of his teaching performance. Specifically, which of four experimental treatments were more and which were less effective in aiding a teacher to modify his student-perceived performance? A detailed description of the four treatments will be given in Chapter III, but to facilitate an initial understanding of the study a brief account is presented here.

As a means of accomplishing the objectives of this study four experimental groups, each employing a prescribed treatment, and one control group were formed. Teachers in treatment group A received written feedback regarding student perceptions of their teaching performances. The feedback data were compiled from an adapted form of the Student Opinion Questionnaire (SOQ), (Appendix A), administered as a pretest to the classes of teachers in this group. The feedback consisted of a graphical description (Appendix B) and a tabular summation (Appendix C) of student perceptions of the teacher's performance. Also included was an informative statement (Appendix D) on possible use of the feedback.

The treatment for teachers in group B was to maintain a record of the various instructional methodologies they employed in their teaching. A checklist log (Appendix E) of educational methodologies, listing both instructional devices and teaching methods, was provided for this purpose. At the conclusion of each day teachers in this group were to indicate, by means of the log, which devices and methods they had employed in their teaching that day. To aid the teacher in making consistent and mutually exclusive checks of the
devices and methods employed, operational descriptions of various methodologies were included with the log. The functional use of the log was thoroughly explained to each teacher prior to the commencement of the experiment.

Teachers in treatment group C met individually with a consultant for approximately one hour each week to make plans for the employment of multiple educational methodologies in their teaching. The consultant was skilled in the use of a variety of methodologies and planned teaching strategy with the teacher. Paramount attention was given to the use of a variety of instructional methodologies, both from day to day and also during any given day. The role of the consultant was most frequently to suggest the available methods for presentation and whenever possible to encourage the teacher to try different and varied methodologies in lesson presentations. Participants in this group also maintained the checklist log previously described.

A combination of the treatments used in groups A, B, and C was utilized by teachers in group D.

A fifth group of teachers, group E, was incorporated as a control group.

For ease in reference group A was called the feedback group, group B was designated the log group, group C was named the planning group, group D was labeled the combination group, and group E was termed the control group.

To further clarify the various group "treatments" an outline synopsis of each group is presented here.
Synopses of the five groups:

A. The feedback group

The teacher: Received feedback information from the SOQ pretest.

B. The log group

The teacher: Maintained a checklist log of his teaching methodologies.

C. The planning group

The teacher:
1. Maintained a checklist log of his teaching methodologies.
2. Planned with a consultant for the employment of multiple educational methodologies.

D. The combination group

The teacher:
1. Received feedback information from the SOQ pretest.
2. Maintained a checklist log of his teaching methodologies.
3. Planned with a consultant for the employment of multiple educational methodologies.

E. The control group

The teacher: Experienced no experimental treatment.

The criterion measure for this study was student perception of teacher performance. An adapted form of the Student Opinion Questionnaire (SOQ) was used to measure the effectiveness of the various group treatments in modifying student perceptions of teacher performances. Items in the SOQ are based on the previous extensive
investigations of Bryan.¹ A more detailed account of the viability of student rating of teacher performance generally and of Bryan's instrument in particular is included in Chapters II and III of this report.

Major Questions Which Were Investigated

In an attempt to carry out the objectives of this study it was desirable to investigate several pertinent questions. These questions were focused basically on two major areas: (1) the success of certain experimental group treatments in helping teachers to modify student perceptions of their teaching performances and (2) the relative success of the group treatments in helping teachers to modify student perceptions of their teaching performances. Rationale for the supposition that the employment of multiple instructional methodologies would improve the performance of teachers as perceived by students is presented in the section of Chapter II entitled "Pedagogical Variety in Teaching."

Because student perception of teacher performance was the dependent variable and the employment of multiple instructional methodologies was the independent variable in this study it was reasoned that students' perceptions of the teacher's overall performance

teaching performance as well as students' perceptions of the teacher's variety in teaching should both be investigated with respect to the two major areas cited above. Certainly it would be crucial to the purposes of the study to determine whether or not each of the prescribed treatment groups was effective in helping teachers to modify student perceptions of their overall teaching performances. Of no less importance would be the desirability of investigating a question to ascertain the relative effectiveness of the various group treatments in helping teachers to modify student perceptions of their overall teaching performances. Similarly, the determination of whether or not each of the prescribed treatment groups was effective in helping teachers to modify student perceptions of their "variety in teaching" as well as the relative effectiveness of the groups in helping teachers to modify student perceptions of their "variety in teaching" were deemed important to the study objectives.

After interaction and discussion with several key people involved in the study, but most importantly with the committee members supervising this project, an exploration of the following four major questions was considered worthwhile and feasible in focusing the investigation on the study objectives:

1. Did each of the groups show a significant difference in overall student perception of teacher performance from pre- to posttest measure?

   It was anticipated that the research effort would provide evidence to show whether in fact a teacher could modify his overall
student-perceived performance by participating in the prescribed treatments.

2. Was there a significant difference between groups on the overall posttest measure?

Due to the fact that the prescribed treatments varied both in nature and degree it was expected that this investigation would determine the relative effectiveness of the various treatments in modifying the teacher's performance as perceived by his students.

3. Did each of the groups show a significant difference in student perception of the teachers' "variety in teaching" (item nine, SQQ) from pre- to posttest measure?

In that the prime independent variable was the use of a variety of instructional methodologies, the study was designed to evaluate the extent to which students perceived a difference in this aspect of the teacher's performance.

4. Was there a significant difference between groups for "variety in teaching" (item nine, SQQ) on the posttest measure?

It was anticipated that the research effort would give indication of the relative effectiveness of the treatments in modifying the students' perceptions of the teachers' variety in instructional methodology.

Definition of Terms

The terms that are used frequently throughout the report are defined for purposes of this experiment as follows:
Average and below average perceived teachers are those teachers whose overall mean scores on the Student Opinion Questionnaire pretest were less than the mean score of the population from which the teachers came. Furthermore, these teachers' pretest scores on item nine of the SOQ, "variety in teaching," were less than the mean score of the population from which the teachers came. These teachers were randomly assigned to the experimental treatments.

The checklist log (Appendix E) is the form listing the various educational methodologies available for teachers. Operational definitions of the methodologies together with spaces for checking dates of use are a part of this form. It was maintained by teachers in experimental groups B, C, and D.

Educational methodology refers to both non-human instructional devices such as projectors, record players, and tape recorders; and also human teaching methods such as lecture, discussion, and role playing.¹

Feedback (Appendixes B, C, and D) refers to the written information provided for the teacher relative to student responses on the questionnaire.

Field experiment is the term used to describe "a research effort in a realistic situation in which independent variables are manipulated under as carefully controlled conditions as possible."^1

An image profile (Appendixes B and C) is the composite of all written feedback.

The term instructional methodology is used interchangeably with educational methodology.

Teaching methods are "patterns of teacher behavior that are recurrent, applicable to various subject matters, characteristic of more than one teacher, and relevant to learning."^2

Teaching strategy is the selection of appropriate educational methodologies for instruction.

A self-contained classroom, for purposes of this study, is a learning situation where one teacher is responsible for more than fifty percent of the learning activities for those students assigned to him. The setting is primarily confined to a single classroom.

Importance of the Study

This study should be of importance to those concerned with the process of education. For those educators who have as their primary

---


2op. cit.
responsibility an influencing of the pedagogical dimension of teaching the study should prove particularly important. In this time of increased accountability it would be desirable to know whether, and to what extent, teachers could modify their perceived effectiveness. The effort described in this report was an attempt to assist teachers in modifying their student-perceived performances by the relatively modest means of adapting and refining the teachers' educational methods and strategies. Should the findings of this study indicate that teachers can modify student perceptions of their teaching performances by employing a variety of instructional methodologies, there would be implications for the development and modification of both pre-service and in-service teacher education programs. Such data could render feasible a more meaningful effort by department and grade group chairmen, principals, and curriculum supervisors to aid their colleagues in the pedagogical aspects of teaching. Based on the results of this study, teachers, as concerned professionals, might be inclined to modify teaching behavior with respect to their instructional methodologies.

Specifically, this field experiment dealt only with first, second, and third-year upper elementary teachers in suburban upper elementary classrooms. Teachers participating in the experimental phase of the study were those who scored average or below average on a written evaluation eliciting student perceptions of teacher performance. The populations to which generalizations could be made, then, would be to those populations similar in nature to the one described here and in the "setting" of Chapter III.
Limitations of the Study

This field experiment had limitations which are readily apparent. One limitation was the fact that, although fifty-four teachers participated in the initial phase of the project, only fifteen teachers were involved in the experimental phase. The practical considerations of time, personnel, and financial resources did not make feasible the expansion of this number. Secondly, the duration of the experiment was for approximately eight weeks during the second semester of the 1969-70 school year. It is possible that this period of time was not sufficient to maximally change teacher performances and subsequent student perceptions or that some long-term effects of the experiment were not evident when the SOQ posttest had been administered. Finally, the effects of the planning sessions were dependent on the efforts of a single consultant and his success in establishing the necessary rapport with individual teachers to influence their use of instructional methodologies.

Assumptions

There were two basic assumptions underlying this investigation; the first pertained to the data gathering instrument while the second related to students' perceptions.

In the absence of definite and agreed on criteria of teacher competence it is difficult to establish direct proof that an instrument designed to reveal the quality of a teacher's performance
actually measures what it purports to measure.\textsuperscript{1} Bryan's instrument, however, has been used extensively for both research and evaluation purposes and is composed of items which are termed "essential characteristics of effective teachers."\textsuperscript{2} For the present study Bryan's instrument was adapted only for the purpose of increasing its understandability at the elementary level. The adaptation was acceptable to teachers in the field, a panel of doctoral students, and the committee supervising this project. Based on the above considerations the adapted instrument was assumed to be appropriate as a measure of student perception of teacher performance.

A second assumption of the study was that there is a meaningful relationship between the students' real perceptions of their teacher's performance and the students' perceptions as elicited by the criterion measures of teacher performance included in the adapted form of the SOQ.

Organization of the Report

Beyond the present chapter, this report will be presented in the following manner: Chapter II, Review of Related Literature and Research, will include an examination of literature and studies pertinent to this experiment. Reviewed under major headings will be


the topics entitled pedagogical variety in teaching, viability of student rating of teachers, student opinion feedback for teachers, motivation of log activity, and function of the consultant.

Following a brief review of the problem and a listing of the questions to be investigated, Chapter III will focus on the research design of the study, the setting, and the procedures used in conducting the field experiment. Included will be the methods used in instrument selection, sample selection, and administration of the instrument. Also presented will be a detailed description of the various group treatments. An account of the procedures used in collecting and organizing the data will conclude Chapter III, Research Design, Setting, and Procedures.

Chapter IV, Presentation and Analysis of the Data, will contain the results and an analysis of the data collected in accordance with the purposes of this investigation.

Presented in Chapter V, Summary, Conclusions, and Recommendations, will be a review of the problem and procedures used, a summary of the findings of the investigation, conclusions, and recommendations.

Summary

This chapter has included a statement of the problem, the objectives of the study, and the major questions which were investigated to provide data on the problem under consideration. Also contained here were the definitions of frequently used terms together with the importance, limitations, and assumptions of the study. Finally, this beginning chapter included an overview of the entire report.
CHAPTER II

REVIEW OF RELATED LITERATURE AND RESEARCH

Information garnered from related literature and research studies pertinent to the investigation under consideration will be presented under the following major headings: (1) Pedagogical Variety in Teaching; (2) Viability of Student Rating of Teacher Performance; (3) Student Opinion Feedback for Teachers; (4) Motivation of Log Activity and (5) Function of the Consultant.

Pedagogical Variety in Teaching

There are three generally accepted dimensions to the teacher's role: (1) his identification with students, (2) his knowledge of subject matter, and (3) his pedagogy or instructional methods skills. Anderson\(^1\) depicts this situation with the diagram shown in Figure 1. Intentionally, instructional methods (pedagogy) occupies the mathematically most significant (Pythagorean Theorem) part of the triangle, for Anderson purports it to be the dimension on which the teacher's professional reputation depends. Colman\(^2\) likewise recognizes these three dimensions but focuses primarily on the teacher's...


use of varied instructional methodologies.

Beck recently conducted a research study with over two thousand sixth-grade students to determine, via a one hundred-item questionnaire, their perceptions of effective teachers. Pupil perceptions were described along five dimensions of teacher merit: affective, cognitive, disciplinary, motivational, and innovative (flexibility and use of various teaching methods). He concluded after factor analysis that the sixth-grade students tended to perceive the

---

effective teacher as a "warm, friendly, and supportive person who communicates clearly, motivates and disciplines pupils effectively, and is flexible in methodology." Using identical statistical methods about five years earlier Gupta\(^1\) had discovered similar factors with high loadings. In another study Sprinthall\(^2\) and his associates focused on "cognitive flexibility" as opposed to "rigidity." They summarized:

"Among teacher educators there is support for the general notion that flexibility or conceptual openness is both a desirable and differentiating quality in teaching. For example, Goodlad has suggested that flexible teaching behavior is most relevant to effective classroom performance ....Translating to the classroom, this (rigidity as opposed to flexibility) would result in the employment of one teaching method until well after its particular contribution has been utilized effectively. The cognitively rigid teacher would miss the cues that another method was needed."

In conclusion the authors of the study supported the basic hypothesis that effective teaching and cognitive flexibility were related.

Most studies to date in this area have attempted to determine what constitutes an effective teacher, but few have attempted to change the students' perceptions of the teacher. Review of the literature uncovered no research that attempted to alter student perceptions of teachers through the use of instructional methodologies. Yet, Gage\(^3\) says that research in instructional methodologies

---


constitutes one of the most important and promising areas in education. He claims it is a neglected and undernourished area of research. Noted educational theorist, Bruner, states that "more than ever before we are concerned with the nature of the educational process...with the techniques and devices that can be used to improve the educational enterprise."

From data received at the Student Reaction Center at Western Michigan University, Bryan has shown that secondary school teachers fall short of student expectations more in the areas of variety of teaching procedures and interest-stimulation than in any other categories. Using Bryan's instrument, the Student Opinion Questionnaire, in a recent study, Lauroesch and his colleagues showed that variety in teaching ranked tenth out of twelve items on mean score. In another study, this time at the elementary level, Koskenniemi reported that "most unsuccessful teachers appeared to lack sensitivity and understanding of children's thinking and attitudes and to exhibit low capacity for effective structuring of instructional activities

---

Furthermore, Amidon and Flanders have found that better teachers showed variety in patterns of teaching behavior while patterns that were similar and consistent characterized the poorer teacher. Grobman asked college students to list what they hoped they would never do in the classroom. One of the prominent answers was the antithesis of variety in teaching.

Following a study focusing directly on teaching methods at the collegiate level, Drayer came to this conclusion:

"It would appear that college students admire effective methods and techniques far more than any other single qualification possessed by their instructors....Although many factors contribute to success as a college teacher, the one outstanding factor which contributes to a successful learning situation is the instructor's knowledge of and skill in methods of teaching. Conversely, lack of this knowledge and skill seems to be the chief factor responsible for failure to achieve a desirable learning atmosphere....Students think the most successful instructors are those who use effective methods."

The importance afforded variety in teaching methodology is not a new dimension in teaching. Historically, Herbert (1776-1841), the famous German educational psychologist, deplored the exclusive use of a single teaching method and instead recommended the employment

---


of a wider range of learning activities. Likewise, Pestalozzi (1746-1827), the Swiss educational reformer, emphasized teaching methodologies that would eliminate the mechanical and routine.

It is not likely that educators will find in the literature anyone purporting the virtues of a single best teaching method. The premise that no method is best in any absolute sense was accepted for this study. Clark and Starr commented as follows:

"Any claim that a single method or approach to the teaching of any subject is the best way to teach that subject has the elements of quackery latent in it. However, the teacher who has mastery over a large store of teaching techniques will find himself ready to provide the tactics necessary for almost any situation."

Wallen and Travers substantiate this position by claiming, on the basis of research in the field, that "teaching methods don't seem to make much difference...there is hardly any direct evidence to favor one method over another." However, they also comment:

"The writers see the great need at the present time for an attempt to design a teaching method which makes as much use as possible of a wide range of learning principles. When this is done, there may be some hope of finding a teaching method which is definitely and markedly superior to others which have not been thus systematically designed. There is a possibility that many different teaching methods might be designed which would make use of many principles."

1Ibid.


4loc. cit., p. 500.
Is it not logical, then, to speculate that an eclectic approach, utilizing many methods with their unique inherent learning principles, might be a justifiable approach? The premise then becomes that, although there appears to be no one best teaching method in any absolute sense, it is equally as certain that some methods are better than others in given situations. Support for this contention comes from Gage who concludes as follows after a survey of research on teaching methods:

"Although positive results remain hard to come by, some can be cited to indicate that, depending on which teaching methods are considered, they can make a difference in educational outcomes."

Furthermore, he says that there are some advantages of certain methods of teaching over others and that if employed on "strategic" occasions, these advantages can outweigh the disadvantages. McKeachie, too, claims that "a conclusion that it does not make any difference which methods are used is clearly unjustified. Rather, recent research suggests that decisions about teaching methods do have important consequences."

The need for a study to incorporate the literature and research cited above was apparent. Envisioned was a study that would not try to pit one method against another or attempt to show the efficacy of

---

1 Gage, "Teacher Methods," op. cit., p. 1447.
2 loc. cit., p. 1456.
a single method of teaching, but rather one that would focus on the use of many educational methodologies already at a teacher's disposal. The study would attempt to determine the resultant effects if teachers would utilize the available methods and media instead of allowing them to go unused.\(^1\) Smith, Krouse, and Atkinson\(^2\) address this point specifically in their discussion of the master teacher:

"The master teacher will be fully aware of all materials and aids at his disposal to further the educational program. The outstanding teacher knows that various techniques must be used in teaching. He will use many of these when they are appropriate to the lesson at hand. The outstanding teacher will use committee work on occasion; he will lecture on occasion and he will use field trips on occasion. In fact, the outstanding teacher uses every device or technique he knows about, but he uses them at various times and for specific purposes....Characteristics that identify a master teacher are also found in teachers of lesser ability; the difference is one of degree."

Such a study as the one described above was the intent of this endeavor. With the use of a variety of instructional methods it was anticipated that improved teaching performance would result. The rationale for this hypothesis was based on the importance afforded the pedagogical dimension by writers in the field and the potential a multiplicity of educational methods gave for curing other classroom ills. This rationale was not unlike that of

\(^1\)Streeter, Edward C., "Teacher Competency and Classroom Use of Educational Media." Audiovisual Instruction, XIV (January 1969), 60.

Gardner\(^1\) who posits the idea that a healthy society should be characterized by pluralism, variety, and alternatives. So a healthy teaching situation should espouse many methods that incorporate varied learning principles.

Viability of Student Rating of Teacher Performance

The viability of student rating of teacher performance has received increasingly favorable comments in recent literature. Historically, student ratings of teachers were first reported in 1923 and achieved considerable attention beginning in the 1930's.\(^2\),\(^3\) Opposition to the practice of student ratings can readily be understood by educators who are familiar with the area of teacher evaluation. In separate articles Amatora\(^4\) and Callahan\(^5\) similarly identify some of the arguments against its use:

---


\(^3\)Amatora, Sister Mary, "Teacher Rating by Younger Pupils." *Journal of Teacher Education*, V (June 1954), 149.

\(^4\)Ibid.

\(^5\)Callahan, Sterling G., "Is Teacher Rating by Students a Sound Practice"? *School and Society*, LXIX (February 1949), 98.
1. Students are too immature to render valid judgments.

2. Students' "halo effect" offsets validity. (Halo is defined by English and English\textsuperscript{1} as "the tendency in making an estimate or rating of one characteristic of a person to be influenced by another characteristic or by one's general impression of that person."

3. Students' responses are tainted with emotional incidents such as grades, teacher attitudes, and discipline.

4. Students' anonymous remarks are irresponsible.

5. Teacher morale is lowered.

Howsam\textsuperscript{2} identified yet another reason for the apparent dichotomy between evidence and practice when he said:

"Despite the favorable evidence, there is widespread resistance to the use of pupil ratings probably arising out of the respective roles of student and teacher in our culture."

Amatora\textsuperscript{3} further clarified the situation:

"The first impulse of the uninitiated is usually one of aversion; whereas those who have had some acquaintance with its procedures, who have studied its pros and cons, who have actually participated in such studies are often favorable and at times even enthusiastic in their endorsement of the practice."

The advantages accruing to student ratings are similarly easily


\textsuperscript{2}Howsam, Robert B., "Teacher Evaluation: Facts and Folklore." The National Elementary Principal, XLIII (November 1963), 16.

\textsuperscript{3}Amatora, op. cit.
identified. Amatora\(^1\) and Callahan\(^2\) again note almost identical arguments of proponents:

1. Students are frank.

2. Students daily occupy strategic positions to see both the good and the bad of a teacher's performance.

3. Students' perceptions of teachers are important to subsequent study.

4. Students provide information, not available through other means, with ease, economy, and convenience.

5. Student criticism is good and helpful in determining pedagogical deficiencies and better teacher personality.

Pupil ratings have been the subject of a considerable amount of research in recent years. With few exceptions the practice has been favorably endorsed. Howsam\(^3\) reports on this research as follows:

"With remarkable consistency the findings have shown that pupils are able to make more valid and reliable ratings of teachers than any other group, including administrators, supervisors, and experts. Teachers in these studies have found the pupil ratings to be fair and accurate."

Earlier, Howsam\(^4\) had found four types of rating scales to be commonly

---

\(^1\)loc. cit., p. 150.

\(^2\)Callahan, op. cit., p. 99.

\(^3\)Howsam, "Teacher Evaluation: Facts and Folklore," op. cit.

used in research on teacher performance. He dismissed self-ratings because they were too biased, peer ratings because there was marginal evidence for colleagues, and supervisor and administrator ratings because of low correlation and bias. In his review of studies How-sam found only student ratings to be favorably endorsed. Hickmott constructed a forty-eight-item instrument for obtaining student opinions of teachers. She reported that correlations for the various groups of fourth, fifth, and sixth-grade students obtained by the chance-half method ranged from .91-.99. The almost unanimous acclaim of teachers regarding the fairness of student responses was supportive of Bryan's earlier work. Amatora did research with pupil ratings in grades four to eight using a five point scale to determine student perception of teacher performance on seven criterion areas. She concluded that elementary pupils are fairly stable in their rating of teachers and show a satisfactory degree of both agreement and discrimination. Split-half reliabilities ranging from .86-.96 were reported on the seven items.

In a monumental attempt to discover criteria for determining teacher merit, McCall conducted a statewide study for the North Carolina State Board of Education. Pupil gain scores on a variety of tests were accepted as being the criterion measure of teacher

---


2loc. cit., p. 44.

3Amatora, op. cit., p. 152.

merit. The judgment of several groups of raters on various teacher traits led him to this conclusion:

"At last we find some professionally competent judges of teaching skill, namely, the teachers' pupils, especially after they have been taught by the teacher for nearly a year. Out of the mouths of children comes more accurate judgment of teachers than that rendered by their peers or supervisors, and, if our criterion is valid, they appear to have a truer idea of what constitutes good teaching than professors of education."

One of the pioneers in research on student rating of teachers has been Remmers with his work on the Purdue Rating Scale for Instruction. This scale has been the tool of research since the 1920's and has led Remmers to conclude that "student evaluation is a useful, convenient, reliable means of self-supervision and self-improvement for the teacher." Major generalizations from researches on this scale follow:

1. Reliability of ratings of teachers by students is a function of the number of raters, in accordance with the Spearman-Brown prophecy formula. If 25 or more student ratings are averaged, they are as reliable as the better educational and mental tests at present available.

2. Grades of students have little if any relationship to their ratings of instructors who assigned the grades.

3. Alumni 10 years after graduation agree very closely (rank order rho = .92) with on-campus students on the relative importance of 10 teacher characteristics.

4. Alumni 10 years after graduation agree substantially (r's ranging from .40 to .68) with on-campus students in their average ratings of the same instructors.

---


2loc. cit., p. 367.

3ibid.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
5. Halo effect, if present in ratings by such instruments as the Purdue Rating Scale for Instruction, is sufficient to raise the intertrait correlations to unity when corrected for unreliability of the ratings. Evidence indicates that students discriminate reliably among different aspects of the teacher's personality and the course.

6. Little if any relationship exists between students' ratings of the teacher and the difficulty of the course.

7. In a given college or university, wide and important departmental differences in teaching effectiveness may exist as judged by student opinion.

8. The sex of student raters bears little or no relationship to their ratings of teachers.

9. The cost in time and money of obtaining student ratings of teachers is low. In fact, it is considerably lower than the cost of administering a typical standardized educational test of some comprehensiveness.

10. Popularity in extraclass activities of the teacher is probably not appreciably related to student ratings of that teacher.

11. Teachers with less than five years' experience tend to be rated lower than teachers with more than eight years' experience.

12. The sex of the teacher is in general unrelated to the ratings received.

13. There is a low but significant positive relationship (r = .20) between the mean objectively measured achievement of an instructor's students (with scholastic ability held constant) and students' ratings of college chemistry teachers.

14. Students are more favorable than instructors to student rating of instructors, but more instructors than students have noticed improvement in their teaching as a result of student ratings."

Miklich recently completed a validation study of the Purdue Rating Scale for Instruction. Educational and Psychological Measurement, XXIX (Winter 1969), 966.
Rating Scale for Instruction and stated:

"We may conclude with some confidence that students can make valid ratings with the Purdue Rating Scale for Instruction or similar kinds of scales."

Davenport\(^1\) likewise concluded that "it can be stated with a fair degree of confidence that pupils are competent to rate teachers, and that their ratings are reliable and valid, and that the ratings of pupils have no deleterious effects on either pupil or teacher morale."

General comments like "research indicates that very real value may be attached to pupil perception of teacher effectiveness as evidence in the evaluation of teachers" from Beck\(^2\) and "students constitute a pool of reliable observers who are in a favorable position to observe changes in the behavior of their teachers since they are present every day" from Oliver\(^3\) are undoubtedly instrumental in the growing trend toward the use of student ratings of teacher effectiveness.\(^4\)

The concepts of rating and evaluation can hardly be discussed apart from the consideration of criteria upon which these judgments


\(^2\)Beck, op. cit.

\(^3\)Oliver, op. cit., p. 5.

are made. Fattu, Mitzel, and Kerlingor among others, posit pessimistic views of past research efforts to yield meaningful, measurable criteria of teacher effectiveness. Yet, the desire to press forward has captivated the efforts of many. The rationale behind such efforts was realistically summarized in an article by the American Association of School Administrators:

"This study is approached with the full realization that there is no absolute certainty anywhere in human affairs. The feelings, beliefs, and psychological reactions of individuals, particularly as they are related to the delicate and intricate process of teaching and learning, cannot be rigidly defined and categorized like inanimate particles of physical matter. But to delay action until such certainty has been achieved would be to delay action forever. Moral obligation rests upon all of us to press forward in the way that all available evidence indicates is more probably true than any other."

Ryans, too, expressed the rationale behind his ambitious endeavor, since called the "single most extensive study of teachers to date":

"While extreme caution should be taken in guarding against an overgeneralized picture of the good or effective teacher, or

---


the opposite exemplified by the inferior or ineffective teacher, the results of a variety of investigations do point to certain recurring descriptions which may have some validity insofar as contemporary culture in the United States is concerned. Certainly the evidence suggests leads and clues which provide starting points for thinking about teacher competencies and for more intensive investigations which open the way for more adequate conceptualizing about teacher performance."

Ryan's study of teacher characteristics has drawn considerable attention and support. Three dimensions or criteria of classroom behavior were identified:

"Pattern X₀: Understanding, friendly vs. aloof, egocentric, restricted teacher behavior.

Pattern Y₀: Responsible, businesslike, systematic vs. evading, unplanned, slipshod teacher behavior.

Pattern Z₀: Stimulating, imaginative, surging or enthusiastic vs. dull, routine teacher behavior."

These three dimensions were identified by surveying the literature, analyzing "critical incidents" in teaching, and identifying the clusters in factor analysis. In a more recent study by Crawford and Bradshaw the effective teacher characteristics determined via three hundred student papers "corresponded closely to those of Ryan in 1960 which presumably represented a consensus of the American Council on Education." Kerlinger similarly states that "there are three principal factors underlying perceptions of desirable traits of

1op. cit., p. 77.


3Kerlinger, op. cit., p. 654.
teachers. These three factors, moreover, resemble Ryans' $X_0 - Y_0 - Z_0$ patterns."

Research literature has in the past divided criteria of effectiveness into presage, process, and product criteria.\(^1\) Presage criteria include factors which teachers bring with them to the classroom. Years of teaching experience, prior preparation, intelligence, and social skills might be such factors. Mitzel\(^2\) commented on the nature of presage criteria:

"In a sense they are pseudo criteria, for their relevance depends on an assumed or conjectural relationship to other criteria, either process or product. Precedent forces their consideration as criteria, since the bulk of research on teacher competence has employed dependent variables which fit into this category."

Soar\(^3\) described process criteria in the following manner:

"Process criteria are aspects of the classroom operation which are deemed worthwhile in their own right, although they may not be directly related to the outcomes of education, the product criteria. These process criteria are most often measures of classroom climate or typical situations involving the social interaction of students and teachers. Other examples would be the extent to which teachers discipline students effectively, maintain rapport with students...."

To this Mitzel\(^4\) added that as process criteria neither student nor teacher behavior should be studied in isolation. "The interaction


\(^2\) Mitzel, op. cit., p. 1484.


\(^4\) op. cit.
between them appears to be the dominant aspect of the whole process of learning." Bryan\(^1\) extended this reasoning when he said:

"One is thinking in terms of process criteria when he concludes that improved student reactions means improved teacher effectiveness. Other things being equal, the teacher who conducts classes that students find challenging and interesting is more effective than the one who conducts classes that bore students. The teacher who gets cooperation in the pursuit of classroom objectives is more effective than one who fails to get students to concentrate on classroom business, and the teacher who is admired and respected is more effective than one who is feared or regarded with contempt."

Product criteria refer to measured growth or gains in students resulting from efforts of the teacher. Fattu,\(^2\) in his manuscript preparation for the American Association of School Administrators, aptly described this criterion area when he concluded:

"Despite fifty years of continued development in the field of educational measurement, satisfactory tests of achievement exist only in a few of the basic skills areas. Adequate measures of social and emotional adjustments, cultural appreciations, or attitudes essential to democratic living are not yet available. Though elaborate statistical and experimental methods have been developed, there is no one who can demonstrate a scientific way of making use of pupil-gain criteria in measuring teacher effectiveness."

Howsam,\(^3\) too, addressed himself to product criteria:

"There is only one fully defensible criterion for judging teacher effectiveness; the ultimate criterion is result. Since it is not presently feasible to rely on the ultimate criterion of effectiveness, it becomes necessary to attempt to develop intermediate or proximate criteria. These are

\(^1\)Bryan, *A Service Designed to Improve the High School Teacher's Image With Students*, op. cit., p. 11.

\(^2\)American Association of School Administrators, op. cit., p. 19.

\(^3\)Howsam, "Teacher Evaluation: Facts and Folklore," op. cit., p. 15.
the ones that can be demonstrated or assumed to be closely related to ultimate criteria."

Upon concluding that of the three major criteria only process criteria were presently of practical use in the teacher's effort for improved effectiveness, Bryan launched a major effort to provide teachers with student opinions of teacher performance. These efforts culminated in the development of the Student Opinion Questionnaire (SOQ) and its subsequent distribution through the Student Reaction Center at Western Michigan University. The questionnaire has been revised and is now called the Teacher Image Questionnaire, while the center has expanded its scope and has been renamed the Educator Feedback Center.

Lauroesch claims that over a period of thirty years Bryan has developed the most reliable procedure for soliciting student opinion of teacher performance. Bryan's twelve-item instrument has been used nationwide in both research and in-service efforts. It has been shown to be reliable, valid, and reasonably free from halo effect. The questionnaire items represent "essential characteristics of effective teachers" and were selected only after a careful study of appropriate research investigations. Of signal importance in this regard was the previously mentioned study of Ryans.

1Bryan, A Service Designed to Improve the High School Teacher's Image With Students, op. cit., p. 11.
2Lauroesch, op. cit., p. 5.
3op. cit., pp. 10, 12.
4op. cit., p. 6.
Ryan's description of the effective behaviors of teachers was largely incorporated in the SOQ. With respect to the relation between teacher effectiveness and student reactions, Bryan concluded:

"Many educators believe that improved student reactions means improved teacher effectiveness. The research by Ryan and others led to the conclusion that effective teachers exhibit to a favorable degree, qualities like those specified in the Student Opinion Questionnaire."

Coats has factor analyzed the results of over forty thousand student responses on the SOQ. He found one factor, which he labeled teacher charisma, to account for about sixty percent of the variance. The other forty percent of the variance was independent of the charisma factor and most likely represented a fairly objective measure of student judgment. Coats added that "teacher charisma is probably a function of teacher effectiveness." Alas, Coats had received consistent requests from boards of education prior to the 1970 school year to administer the SOQ for purposes of merit rating. Although the teacher effectiveness criteria dilemma is still bothersome for many educators, the above-mentioned use of this instrument has made more common its designation as an instrument for measuring teacher effectiveness.

---

1Ryan, op. cit., p. 82.


effectiveness.

Many pre- and posttest designed studies such as the one carried out here do not attempt to engage in any absolute evaluation of teacher effectiveness. No attempt was made at comparison with external standards. Rather, the effectiveness of various group treatments was determined by changes in students' pre- and posttest opinions. As such the effort was valid in its solicitation of student perceptions of teacher performance.

Student Opinion Feedback for Teachers

The rationale for hypothesizing changes in teacher behavior based on student opinion feedback has been offered by many writers.\(^1\) Gage, Runkel, and Chatterjee claimed an imbalance was established which prompted the teacher to move in directions of correction. Osgood and Tannenbaum referred to this same phenomenon as incongruity, Festinger as dissonance. Given time following feedback, subjects will adjust their behavior to restore consistency or equilibrium.

Many have espoused the concept of feedback as an aid in improving teacher performance. Not infrequently teachers have engaged of their own volition in attempts to get feedback from students regarding their teaching. In explaining the benefits of such activity Tedesco\(^2\)

---

\(^1\)Daw, R. W. and Gage, N. L., "Effect of Feedback from Teachers to Principals." *Journal of Educational Psychology*, LVIII (June 1967), 181.

\(^2\)Tedesco, Phyliss Reynolds, "An 'A' for Teacher." *The Instructor*, LXIX (October 1959), 89.
said, "I read those papers five times. I learned more about my teaching methods that evening than I would have in a year of course work." Wright and Sherman\(^1\) also agreed:

"Whatever may be the teacher's evaluation of herself or a supervisor's evaluation of her, neither can compare with pupils' perceptions of the teacher as a way to find out what pupils are responding to."

Other attempts at providing teacher feedback have become more systematized such as the efforts at Purdue, Stanford, and Western Michigan Universities. In his doctoral dissertation at Stanford, Aubertine\(^2\) found that participants valued most the feedback easiest to obtain, namely that from students. Ryan\(^3\), studying at the same institution, also found that pre-service teachers perceived student feedback to be more beneficial than that from other sources.

After reviewing material in the field and prior to conducting a study using Bryan's instrument, Lauroesch\(^4\) and his associates concluded that "thirty years of research and development have indicated that student feedback is a useful and reliable means for improving and directing behavioral change in teachers." In a study conducted

\(^1\)Wright, Benjamin and Sherman, Barbara, "Love and Mastery in the Child's Image." School Review, LXXIII (Summer 1965), 89.


\(^3\)Ryan, Kevin A., "The Use of Students' Written Feedback in Changing the Behavior of Beginning Secondary School Teachers." Dissertation Abstracts: The Humanities and Social Sciences, XXVII (January-February 1967), 2089-A.

\(^4\)Lauroesch, op. cit., p. 1.
with a USOE grant Bryan\(^1\) found that "the feedback of information about student reactions can be used by many teachers as a means of improving effectiveness as seen by students." Fifty-seven percent of sixty experimental teachers made a statistically significant gain on one or more questions as compared with twenty-four percent of the fifty-nine non-experimental teachers. Ninety-two percent of the experimental group teachers said the feedback was helpful. Oliver\(^2\) used Bryan's instrument in a recent study and drew these conclusions:

1. Informational feedback from students is effective in changing teacher behavior.

2. Student feedback is more effective in changing teacher behavior than supervisory feedback.

3. The utilization of student feedback as a means of improving teacher behavior should be used to a greater extent."

McCall's\(^3\) previously cited and comprehensive study in North Carolina indicates that for training teachers in service "a substantial gain could be secured by the simple device of having pupils give a confidential rating of their teachers at the end of the school year."

It is apparent from the literature and research studies that student feedback is beneficial for many teachers in bringing about improved teacher performance as perceived by students. It has consistently been regarded as more valuable than that of supervisors,

\(^1\)Bryan, *Reactions to Teachers by Students, Parents, and Administrators*, op. cit., p. 43.

\(^2\)Oliver, op. cit., pp. 47-8.

\(^3\)McCall, op. cit., p. 37.
reliable, and reasonably free from the influence of "halo." To the extent that rating instruments incorporate items characteristic of effective teachers, the feedback could be expected to improve the effectiveness of many teachers as perceived by students.

Motivation of Log Activity

It was expected that teachers maintaining the checklist log would be motivated as a result of that log to improve their variety in teaching. The rationale for this expectation came largely from the psychology of motivation. Well-documented in the annals of educational and social psychology was the complexity and importance of motivation as a predisposition for certain behavior. From an examination of motivation theory literature three salient features of the log were readily apparent:

1. The log provided participants with the stimulus or motivation to perform in ways commensurate with their own personal expectations for fulfilling a given role.

2. The log provided participants with the stimulus or motivation to perform in ways commensurate with their own perceptions of others' role expectations for them.

3. The log provided an account of past activities as an aid to future desired performance. Each of these features is here expanded in turn as it pertains to the checklist log of this study.

Kretch and Crutchfield\(^1\) describe the state of motivation as "the

most striking feature of the person's experience of self." Additionally, they state:

"The person perceives disturbances and deficiencies with respect to himself and his surroundings. He feels needs and desires, he sets goals and forms intentions, he exercises choice and will. In response to these motivations the person acts."

The mere presence of the log creates an awareness, a knowledge of existing instructional methodologies. Whatever motivations and subsequent actions taken by the teacher at this point are most likely a response to the discrepancy between what he feels his variety in teaching is (the self-real image of his performance with respect to this variety criterion) and what he feels it ought to be (the self-ideal image of his performance with respect to this variety criterion).

Mouly¹ claims that the self-ideal "represents his standards of conduct based upon his interpretations of the role prescriptions relayed to him...." It was intended that the log would be a factor in the teacher's own role prescription with respect to his variety in teaching.

In their discussion of motives Kretch and Crutchfield² identified two basic types of motives, deficiency motives and abundancy motives. Among others the former variety was characterized by "need to remove deficits" and might be described as tension-reduction behavior. The latter variety is characterized by "desires to know,

²op. cit., p. 278.
understand, learn, create, achieve..." and might be accompanied by tension increase. Additionally, Kretch and Crutchfield\textsuperscript{1} state:

"Every individual is characterized by both deficiency and abundancy motivations. But the relative weighting of the two varies widely among individuals. Some persons are mainly dominated by their needs, with desires impoverished; in others needs may be subsidiary and desires have full range. Almost every activity we find people engaged in can express either deficiency motivation or abundancy motivation."

It is probable, then, that the external stimulus provided by the log motivated participants to act in ways to alleviate deficiencies. In accord with the discussion of Kretch and Crutchfield there could have been varying motives underlying the action, but undoubtedly any response would have been directed toward equilibrium—a "homeostasis" of the social environment—in an attempt by the individual to reduce the imbalance, deficiency, or incongruity precipitated by the log.

In addition to the above line of reasoning it is entirely possible that the participants received personal reinforcement merely from maintaining the checklist log. According to this explanation, based on behavioral psychology, participants would have been motivated to repeat the behavior (variety in methodologies) for which reinforcement was possible through checking the log.

A second feature of the log which was identified previously dealt with the stimulus provided by it to perform according to the expectations of others. To the extent that participants felt an achievement-evaluation type of motivation the log provided a form of surveillance of their instructional methodologies. In effect, the log was an external authority imposed on the teachers with respect to...
their instructional methodologies. It was likely that the factor of external authority was not a dominating force because the researcher did not represent a power figure or occupy an authority position with respect to the teacher participants. However, the motivation to please the researcher might well have been similar to the action of pleasing a superior.

The third and final feature of the log was the opportunity it provided to account in a systematic way for methodologies employed and to use this record as a guide for determining future methods. The log was suggestive in that it listed many possible methods from which to choose in planning teaching strategy. In this respect it was similar to the rationale underlying a categorization of questions featured in a book by Sanders.\(^1\) He claimed that "after a teacher studies the taxonomy he is likely to offer his students a greater variety of intellectual experiences than he did before." The comparable use of the log fostered a cognition of what was available and what had been used. Festinger\(^2\) refers to the disparity between cognitions which exist simultaneously for a person as cognitive dissonance. His reference to cognitive dissonance as a motivating state relates this final feature of the log to the first:


"I wish to hypothesize that the existence of cognitive dissonance is comparable to any other need state. Just as hunger is motivating, cognitive dissonance is motivating. Cognitive dissonance will give rise to activity oriented toward or eliminating the dissonance."\(^1\)

The entire situation posited here is not unlike that of teachers requiring students to keep a log of their activities, principals requiring teachers to maintain a plan book, or employers requiring employees to account for their endeavors via a time log. In all cases the three-fold features of the log exist as described above. To make use of the log in relation to these three features was the intent behind the introduction of this variable into the field experiment.

Function of the Consultant

Traditionally the role of the supervisor in education has been considered a line position; supervision has been assumed by a person occupying an administrative post. The nature of this supervision was largely inspection. Elsbree, McNally, and Wynn\(^2\) described as follows this situation:

"In its earliest form, supervision meant inspecting the work of the teacher....Inspection was not an attempt to help teachers improve instruction. It was designed to determine whether or not teachers did what they were supposed to do, and if they didn't, to replace them with teachers who would."

The supervision of instruction next passed through a period in the

\(^1\)loc. cit., p. 171.

early decades of this century of "teaching the teachers how to do it." Explicit in this concept of supervision was the idea that the supervisor knew best how to teach and therefore should attempt to indoctrinate the staff in his superior methods. Desirable as it may be for the supervisor to possess this superiority it is very doubtful that the intended outcome was often realized. With the increased quantity and quality of pre-service preparation teachers generally bring a degree of expertise to the educational scene. A recognition of this fact has prompted a shift to the idea of supervision as democratic leadership whereby the supervisor and teacher together explore the means of improving the total teaching-learning situation. Mason likewise concurs that supervision is assistance rather than inspection.

Although the remnants of past supervisory practice still linger too frequently, it has become increasingly popular for the supervisor to be looked upon as a helper. This role of helper or consultant is described by Goldberg:

"As a catalyst the supervisor causes a situation to come about--to be the prime mover, yet to create a feeling in the teacher that this was an effort on his part."

Elsbree, McNally, and Wynn state unequivocally that "the task of instructional improvement in the local school is inescapably the

---

1loc. cit., p. 142.

2Mason, Barbara T., "Supervisor or Curriculum Specialist." Educational Leadership, XXVII (January 1970), 401.


4Elsbree, McNally, and Wynn, op. cit., p. 167.
responsibility of the principal." Although it is generally accepted that this is the principal's primary function, the administrivia of his position too frequently mitigate against its top priority. Many school systems now have special supervisory services to handle or share this responsibility with the principal.

The objective behind supervision, in its broadest sense, is the process of improving instruction.\(^1\) Sometimes the principal alone functions in this capacity; in other cases personnel with various titles assume this function. These instructional specialists might be called elementary school supervisors, coordinators, consultants, subject matter specialists, curriculum coordinators, or grade group chairman, to mention a few. In almost every instance they desire to be looked upon as "service" personnel with staff rather than line responsibilities. Their emphasis is upon in-service education and assistance for the classroom teacher.\(^2,3\) With respect to the consultant's responsibility Spain, Drummond, and Goodlad\(^4\) conclude that "the focus is upon teacher improvement, with children's betterment a hoped-for outcome, rather than upon direct contact with individual

---


\(^3\) Mason, op. cit., p. 403.

\(^4\) op. cit.
children. In a research study at the sixth-grade level, Ginther\(^1\) found evidence to support the theory that consultants working with teachers rather than students produced better gain scores in student achievement. This indicated a greater impact by consultants if their efforts were with teachers rather than with students.

That there is disagreement over whether supervision should come from a line or staff position is well known.\(^2\) The distinction between the role of staff and line personnel can perhaps best be illustrated with this statement regarding the responsibility of the line administrator:

"The principal's responsibility is to help weak teachers who are not on tenure as much as he can. If after two or three years of working with them, he is convinced that they are unlikely to become reasonably satisfactory teachers, he will have to dismiss them."\(^3\)

Because the staff personnel are not ultimately responsible for dismissal procedures it is perhaps easier for them to divorce themselves from the inspection role of long ago and assume the more contemporary role of dynamic democratic leadership in working with teachers to improve the total teaching-learning situation. Mason\(^4\) is more adamant in her support of the instructional specialist:

---

\(^1\)Ginther, John R., "Achievement in Sixth Grade Science Associated With Two Instructional Roles of Science Consultants." *Journal of Educational Research*, LVII (September 1963), 30.

\(^2\)Heald, op. cit.

\(^3\)Elsbree, McNally, and Wynn, op. cit., p. 167.

\(^4\)Mason, op. cit., p. 403.
"The professional personnel who work with teachers in the area of curriculum should be called curriculum specialists because of the negative connotations of the word supervisor .... When teachers are held accountable for the results rather than the methods, they will welcome, even seek out, those specialists who they think can help them secure the best results. I think they will call first for consultants, curriculum specialists, anyone with a title that does not convey the idea of inspection."

Judging from the prevalence of the practice of engaging the services of consultative personnel who function in a staff position, their efforts must be considered beneficial. Based on the literature¹,²,³ their role might best be considered complementary to that of the building principal, rather than pre-empting him of the opportunity to function in that capacity. Although the literature is replete with the favorable results of in-service education generally,⁴ no direct research evidence was uncovered which would document the improvement of teacher performance through the efforts of the staff specialist.


²Elsbree, McNally, and Wynn, op. cit., 113.


CHAPTER III

RESEARCH DESIGN, SETTING, AND PROCEDURES

This chapter of the report will focus on the research design of the study, the setting, and the procedures used in conducting the field experiment. Included will be the methods used in instrument selection, sample selection, and the administration of the instrument. Also presented will be a detailed description of the various group treatments. An account of the procedures used in collecting and organizing the data will conclude the chapter.

Review of the Problem

The primary purpose of this study was to determine if teachers could modify student perceptions of their teaching performances by employing multiple educational methodologies in their teaching. A secondary objective was to ascertain the relative effectiveness of four experimental treatments in helping a teacher to modify student perceptions of his teaching performance.

Five groups of teachers, employing prescribed experimental treatments, were formed in an attempt to gather data pertinent to the objectives of this investigation. Teachers in experimental group A, referred to as the feedback group, received written feedback from students regarding student perceptions of their teaching performances. The treatment for teachers in group B, the log group, was to maintain a checklist log of the various instructional methodologies they
employed in their teaching. Teachers in group C, known as the planning group, met weekly with a consultant to make plans for the employment of multiple educational methodologies in their teaching. Participants in this group also maintained the checklist log. A combination of the treatments used in groups A, B, and C was utilized by teachers in group D. In addition, a fifth group of teachers, group E, was used as a control group. An overview of these five groups is shown in Table 1.

Table 1
Overview of the Five Treatment Groups

<table>
<thead>
<tr>
<th>Groups</th>
<th>Treatment</th>
<th>Number of Teachers</th>
<th>Number of Students</th>
<th>Grades</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>F</td>
<td>3</td>
<td>77</td>
<td>4, 4, 6</td>
</tr>
<tr>
<td>Group B</td>
<td>L</td>
<td>3</td>
<td>77</td>
<td>4, 4, 6</td>
</tr>
<tr>
<td>Group C</td>
<td>LP</td>
<td>3</td>
<td>72</td>
<td>4, 5, 6</td>
</tr>
<tr>
<td>Group D</td>
<td>FLP</td>
<td>3</td>
<td>67</td>
<td>4, 6, 6</td>
</tr>
<tr>
<td>Group E</td>
<td>C</td>
<td>3</td>
<td>77</td>
<td>4, 5, 6</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>15</td>
<td>370</td>
<td></td>
</tr>
</tbody>
</table>

*F—Feedback
*L—Log
*P—Planning
*C—Control

The criterion measure for obtaining the data was an adapted form of Bryan's Student Opinion Questionnaire (Appendix A).
Major Questions Which Were Investigated

Formulated below are the major questions which were investigated in an attempt to carry out the objectives of this study:

1. Did each of the groups show a significant difference in overall student perception of teacher performance from pre- to posttest measure?

2. Was there a significant difference between groups on the overall posttest measure?

3. Did each of the groups show a significant difference in student perception of the teachers' "variety in teaching" (item nine, SOQ) from pre- to posttest measure?

4. Was there a significant difference between groups for "variety in teaching" (item nine, SOQ) on the posttest measure?

Specific Questions Which Were Investigated

Emanating from the four major questions included in the preceding section were thirty component questions which are here enumerated. Questions 1a through 1e pertain to major question 1 while questions 2a through 2j refer to major question 2. Likewise questions 3a through 3e apply to major question 3 and questions 4a through 4j relate to major question 4.

1a: Was the overall student perception of teacher performance for teachers in group A (written feedback) more favorable on the posttest measure than on the pretest measure?

1b: Was the overall student perception of teacher performance
for teachers in group B (log maintenance) more favorable on the posttest measure than on the pretest measure?

1c: Was the overall student perception of teacher performance for teachers in group C (employment of multiple educational methodologies and log maintenance) more favorable on the posttest measure than on the pretest measure?

1d: Was the overall student perception of teacher performance for teachers in group D (employment of multiple educational methodologies, log maintenance, and written feedback) more favorable on the posttest measure than on the pretest measure?

1e: Was the overall student perception of teacher performance for teachers in group E (control) similar on the pre- and posttest measures?

2a: Was the written feedback of student perception of teacher performance (group A) more effective than no treatment at all in helping a teacher to modify student perception of his teaching performance?

2b: Was the maintenance of the checklist log of various educational methodologies (group B) more effective than no treatment at all in helping a teacher to modify student perception of his teaching performance?

2c: Was the employment of multiple educational methodologies and log maintenance (group C) more effective than no treatment at all in helping a teacher to modify student perception of his teaching performance?

2d: Was the employment of a combination of feedback, log
maintenance, and multiple educational methodologies (group D) more effective than no treatment at all in helping a teacher to modify student perception of his teaching performance?

2e: Was the maintenance of a checklist log of various educational methodologies (group B) more effective than written feedback alone (group A) in helping a teacher to modify student perception of his teaching performance?

2f: Was the employment of multiple educational methodologies supplemented with log maintenance (group C) more effective than written feedback alone (group A) in helping a teacher to modify student perception of his teaching performance?

2g: Was the employment of multiple educational methodologies supplemented with log maintenance and written feedback of student perception of teacher performance (group D) more effective than written feedback alone (group A) in helping a teacher to modify student perception of his teaching performance?

2h: Was the employment of multiple educational methodologies supplemented with log maintenance (group C) more effective than the maintenance of a checklist log alone (group B) in helping a teacher to modify student perception of his teaching performance?

2i: Was the employment of multiple educational methodologies supplemented with log maintenance and written feedback of student perception of teacher performance (group D) more effective than the maintenance of a checklist log alone (group B) in helping a teacher to modify student perception of his teaching performance?

2j: Was the employment of multiple educational methodologies
supplemented with log maintenance and written feedback of student perception of teacher performance (group D) more effective than the employment of multiple educational methodologies and log maintenance alone (group C) in helping a teacher to modify student perception of his teaching performance?

3a: Was student perception of teacher performance regarding "variety in teaching" for teachers in group A (written feedback) more favorable on the posttest measure than on the pretest measure?

3b: Was student perception of teacher performance regarding "variety in teaching" for teachers in group B (log maintenance) more favorable on the posttest measure than on the pretest measure?

3c: Was student perception of teacher performance regarding "variety in teaching" for teachers in group C (employment of multiple educational methodologies and log maintenance) more favorable on the posttest measure than on the pretest measure?

3d: Was student perception of teacher performance regarding "variety in teaching" for teachers in group D (employment of multiple educational methodologies, log maintenance, and written feedback) more favorable on the posttest measure than on the pretest measure?

3e: Was student perception of teacher performance regarding "variety in teaching" for teachers in group E (control) similar on the pre- and posttest measures?

4a: Was the written feedback of student perception of teacher performance (group A) more effective than no treatment at all in helping a teacher to modify student perception of his "variety in teaching"?
4b: Was the maintenance of the checklist log of various educational methodologies (group B) more effective than no treatment at all in helping a teacher to modify student perception of his "variety in teaching"?

4c: Was the employment of multiple educational methodologies and log maintenance (group C) more effective than no treatment at all in helping a teacher to modify student perception of his "variety in teaching"?

4d: Was the employment of a combination of feedback, log maintenance, and multiple educational methodologies (group D) more effective than no treatment at all in helping a teacher to modify student perception of his "variety in teaching"?

4e: Was the maintenance of a checklist log of various educational methodologies (group B) more effective than written feedback alone (group A) in helping a teacher to modify student perception of his "variety in teaching"?

4f: Was the employment of multiple educational methodologies supplemented with log maintenance (group C) more effective than written feedback alone (group A) in helping a teacher to modify student perception of his "variety in teaching"?

4g: Was the employment of multiple educational methodologies supplemented with log maintenance and written feedback of student perception of teacher performance (group D) more effective than written feedback alone (group A) in helping a teacher to modify student perception of his "variety in teaching"?

4h: Was the employment of multiple educational methodologies
supplemented with log maintenance (group C) more effective than the maintenance of a checklist log alone (group B) in helping a teacher to modify student perception of his "variety in teaching"?

4j: Was the employment of multiple educational methodologies supplemented with log maintenance and written feedback of student perception of teacher performance (group D) more effective than the maintenance of a checklist log alone (group B) in helping a teacher to modify student perception of his "variety in teaching"?

4j: Was the employment of multiple educational methodologies supplemented with log maintenance and written feedback of student perception of teacher performance (group D) more effective than the employment of multiple educational methodologies and log maintenance alone (group C) in helping a teacher to modify student perception of his "variety in teaching"?

Design of the Study

The design of this study is commonly called a "true experimental design."\(^1\) It is labeled "true" because the researcher was able to randomly assign subjects to treatment groups. Because the investigator was able to manipulate an independent variable while measuring the concomitant variation on a dependent variable this investigation can aptly be designated as an "experiment." The pretest--treatment--posttest sequence with an accompanying control group as used in this experiment

is called the "classical design" of research.¹ The research paradigm for the study is shown in Figure 2. Letters "A", "B", "C", "D", and "E" designate the various groups with their accompanying experimental treatments. Representing pre- and posttest scores are the letters "X" and "Y" respectively.

\[
\begin{align*}
\text{A} & \quad \text{X} \quad \text{Y} \\
\text{B} & \quad \text{X} \quad \text{Y} \\
\text{C} & \quad \text{X} \quad \text{Y} \\
\text{D} & \quad \text{X} \quad \text{Y} \\
\text{E} & \quad \text{X} \quad \text{Y}
\end{align*}
\]

Figure 2

Research Paradigm

The independent variable was the type of treatment provided the teachers in each of the groups. The criterion measure employed to collect data on the dependent variable was the score on the written measure of student perception of teacher performance, the Student Opinion Questionnaire. Three teachers were randomly assigned by lottery to each of the groups. Thus, there were approximately seventy-five subjects for each treatment group (three classes of approximately twenty-five students each).

¹loc. cit., p. 309.
The design and rationale of the experiment necessitated the selection of teachers for the experimental phase who were capable of showing improvement on the student perception scores of their teaching performance. Much of the rationale for this decision was based on the previously cited study of Ryan\(^1\) where he reported no significant change because the students' high initial perception of the teacher's performance pre-empted that possibility. Hence, in addition to providing a necessary component in the change score on the basis of pre-posttest data, the pretest measure was used as a means for selecting teachers with improvement possibilities. The design called for fifteen such teachers, three each for the five groups.

**Selection of the Instruments**

The means for gathering the data pertinent to the experiment became a crucial factor in the study. It was necessary to be able to measure the effect, if any, of the experimental treatment on the teacher's performance. To assess this effect it was desirable to get an initial indication of teacher performance as well as a final determination. This would provide a measurement of the net change in teacher performance during the course of the experiment.

Several possibilities existed for obtaining this pre-post information on the performances of the teachers. In addition to the selection of an appropriate data gathering instrument it was necessary to simultaneously consider by whom the evaluation could best be made.

\(^1\)Ryan, op. cit., p. 2039-A.
In his review of research studies on the topic of teacher evaluation, Howsam\(^1\) was most favorably impressed with classroom observation studies. Use of a teacher behavior classification system such as Flanders,\(^2\) the critical incidents technique, or an evaluation device similar in nature to the common and widely used school district evaluation instruments were all possibilities.\(^3\) With respect to a classroom observer(s) there were two fundamental alternatives. One possibility was to use an observer(s) foreign to the class; a school district or building administrator(s), or an observer(s) selected by the researcher were all feasible in this regard. The other alternative was to have teacher performance evaluated by observer(s) regularly witnessing the teacher in action. These observers were, of course, the students. Available resources and the encouraging literature previously cited in Chapter II of this report precipitated a pursuit of student evaluation of teacher performance.

Initially, a careful examination of currently used teacher evaluation forms in several public school systems was made. The items for evaluation corresponded well with the characteristics of master


\(^3\)op. cit., p. 29.
teachers in The Educator's Encyclopedia and also a list of teacher expectations by Davenport. Following this examination, an analysis of the measuring devices available in Buros' Mental Measurement Yearbook, and a survey of the literature cited in Chapter II, the Student Opinion Questionnaire (Appendix A) from the Educator Feedback Center at Western Michigan University was considered an appropriate data gathering instrument to be adapted for use by upper elementary students. The SOQ was designed to elicit student perceptions of teachers on twelve criterion measures of teacher performance. These measures together with their reliability coefficients are shown in Table 2. The reliability coefficients on the various items as reported by Bryan ranged from .77 - .95. These coefficients were correlated by taking the averages of student responses from chance halves of the classes of fifty randomly selected secondary school teachers. Furthermore, studies show consistent test-retest stability of the teacher's image unless well directed efforts to change it have been expended. The objectivity of the instrument should yield

---


4Bryan, Some Observations Concerning Written Student Reactions to High School Teachers, op. cit., p. 3.

a perfect scorer reliability of 1.00.

Table 2
Reliability of Items on Bryan's Student Opinion Questionnaire

<table>
<thead>
<tr>
<th>Item No.</th>
<th>Item</th>
<th>Reliability Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Knowledge of Subject</td>
<td>.87</td>
</tr>
<tr>
<td>2</td>
<td>Clarity of Explanations</td>
<td>.82</td>
</tr>
<tr>
<td>3</td>
<td>Fairness</td>
<td>.84</td>
</tr>
<tr>
<td>4</td>
<td>Control</td>
<td>.95</td>
</tr>
<tr>
<td>5</td>
<td>Attitude Toward Students</td>
<td>.88</td>
</tr>
<tr>
<td>6</td>
<td>Ability to Stimulate Interest</td>
<td>.87</td>
</tr>
<tr>
<td>7</td>
<td>Attitude Toward Subject</td>
<td>.90</td>
</tr>
<tr>
<td>8</td>
<td>Attitude Toward Student Opinions</td>
<td>.86</td>
</tr>
<tr>
<td>9</td>
<td>Variety in Teaching Procedures</td>
<td>.91</td>
</tr>
<tr>
<td>10</td>
<td>Encouragement of Student Participation</td>
<td>.77</td>
</tr>
<tr>
<td>11</td>
<td>Sense of Humor</td>
<td>.91</td>
</tr>
<tr>
<td>12</td>
<td>Planning and Preparation</td>
<td>.90</td>
</tr>
</tbody>
</table>

The face validity of the instrument was deemed acceptable in that the items were "essential characteristics of effective teachers."¹ According to Bryan the items were selected only after a careful study of many appropriate research reports of qualities possessed by effective teachers.

¹Bryan, A Service Designed to Improve the High School Teacher's Image With Students, op. cit., p. 12.
teachers. Of signal importance in this regard was the study conducted by Ryans.\(^1\) Continued examination and revision of this instrument over a period of several years has maintained this content validity. Analyses of the results of the instrument have shown variability in student response between teachers and discriminatory power among test items. Clark\(^2\) was convinced of the viability of the instrument when he concluded that "Bryan's Student Opinion Questionnaire (SOQ) is undoubtedly the most reliable and useful instrument that exists for measuring students' opinions of their teachers."

This instrument, however, was designed for use in grades 7-12 and consequently required an adaptation for use by upper elementary students. Permission from the Educator Feedback Center at Western Michigan University was sought for this adaptation and permission was granted. This endeavor was based primarily on Bryan's original instrument, but also gave consideration to the revisions made by Coats, Bryan's successor as director of the Educator Feedback Center, and Tobin. In consultation with several elementary school teachers the instrument was adapted for upper elementary use. Permission was received from the superintendent, building principal, and teachers in one of the participating districts to field test the instrument for understandability at an upper elementary level. Subsequently, twelve

---

\(^1\)Ryans, op. cit., pp. 1-416.

randomly selected fourth- and fifth-year students from a non-graded elementary school were employed for this part of the study. Each of these students reported individually to a conference room in their building where an attempt was made to ascertain each student's understanding of the adapted SOQ. Each student was given a copy of the adapted SOQ and asked to follow along as the researcher read each item aloud. At the conclusion of each item the student was asked to tell in his own words the meaning of the question. The student was not asked to answer the question, but rather to state in his own words what he thought the question asked.

Each of the sessions with the twelve students was taped so that the responses could be judged for understandability by a panel of three doctoral students studying in the area of educational leadership. Using a scoring sheet specifically designed for that purpose, the panel of judges unanimously agreed that the students were successfully able to comprehend the questionnaire items.

Some minor changes based on the recommendations of the judges were made to reduce ambiguity in the adapted SOQ. Prior to drafting the final instrument for printing, the suggestions of committee members regarding ambiguity, clarity, and the continuity of question to answer were incorporated in the instrument. To insure appropriate scaling the book entitled Scales for the Measurement of Attitudes by Shaw and Wright¹ was consulted. The final draft was mutually

acceptable to the committee, the elementary teachers whose aid was solicited, the panel of judges, and the researcher.

A second instrument, the checklist log of educational methodologies (Appendix E), was used by teachers in experimental groups B, C, and D. It was refined specifically for this study. The initial checklist was developed in cooperation with a group of prospective teachers in an educational methods class taught by the researcher at Hope College in 1969. Prior to completing this form an examination was made of several general education methods textbooks, the Encyclopedia of Educational Research, the Handbook of Research on Teaching, The Educator's Encyclopedia, and the Dictionary of Education. This was done in an effort to be exhaustive with respect to the inclusion of major methodologies and also to aid in operationally describing those methods which might be ambiguous. Minor changes in content and form were made after consulting the literature.

The log was field tested by the same group of fourth- and fifth-grade teachers whose services were rendered in the initial testing of the adapted SOQ. They found the form to be functional, a minor imposition on their time, and of considerable potential value for participating teachers. Their suggestions regarding form, as well as those of the committee, were incorporated in the final instrument.

The log was maintained as part of the experimental treatment for teachers in groups C and D. It was the only treatment for teachers in group B. In addition to this function with regard to the independent variable the log served as an indicator of the degree to
which the independent variable was manipulated.

The Setting

This field experiment was implemented in three contiguous suburban school districts in southeastern Michigan. Selected because of similarity in membership, cost per membership pupil, valuation per pupil, and geographic location, the three districts were frequently considered comparable by superintendents within the intermediate school district.¹ Each of the districts had experienced substantial growth during the five years just preceding the study. This growth was due to the not uncommon exodus of white middle class families from an adjacent integrated industrial city. The population of each of the districts was almost exclusively white. Although to be sure there were visible differences in the socio-economic background of the families residing within the various school attendance areas within each of the districts, these differences were believed minimal in comparison with the differences inherent in large districts with accompanying inner-city schools. The research of Tobin² supports this notion that school setting is a factor in student perception of teacher performance. Working with a similar instrument at comparable


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
age levels he found that suburban teachers were viewed more favorably than were inner city teachers. Because the partitioning of subjects based on setting was not a part of the design of this study, this attempt at homogeneity within the setting was considered important in the control of variables.

Procedures

Following committee approval of the dissertation prospectus in early December of 1969 a personal contact was made with the superintendent in each of the prospective districts. The experiment was described in detail and two of the superintendents gave immediate approval for the study. In accord with the designated teacher criteria for the study these two superintendents provided the names of teachers who were in their first, second, or third year of teaching and were assigned to a fourth, fifth, or sixth-grade self-contained classroom. Also provided were the names of the corresponding building principals. Additionally, permission was granted to contact principals and teachers personally. The third superintendent gave his immediate approval, but withheld permission to proceed further pending board approval as dictated by school policy. Following the submission to the board of a prescribed complex written description of the experiment, approval was granted by a designated board committee in early January of 1970. At this point names of teachers meeting the study criteria and also the names of the corresponding building principals were provided by the third school district superintendent.
He, too, gave permission to contact teachers and principals personally. Time was arranged on the agenda at the next principals' meeting in each of the three districts. The study was again explained in detail and the names of teachers meeting the established criteria were verified by the principals. The principals' roles in the experiment were minimal. It was deemed desirable, however, for them to be informed of the nature of the study and to solicit their cooperation in procedural matters. Obviously, the details of the study as explained to the principals were not to be shared with their teaching staffs. The principals aided in the scheduling of individual appointments with the researcher and the teachers. Their instructions were simply to "inform individually each of the eligible teachers in your building that an appointment had been made for them with a researcher to discuss a research project on such and such a date at such and such a time."

The names of fifty-four teachers meeting the established criteria of the study were received from the three districts. The teachers' assignments were in twenty-one different buildings. In an attempt to eliminate any peer group pressure, either to participate or to refrain from participating, individual appointments were arranged through the respective building principals with each of the fifty-four teachers. These appointments were scheduled at the various school buildings both before and after school and also during the teachers' planning periods throughout the school day.

An introductory cover letter (Appendix F) from each of the
district superintendents was presented to each teacher at the time of the initial contacts. The fifty-four prospective participants were individually informed of the following information:

1. That the study dealt with the nature of the instructional process.

2. That the investigator was requesting the teacher's voluntary permission, as a preliminary phase of the study, to personally administer to their class the adapted Student Opinion Questionnaire. The teachers were further informed that the questionnaire asked the students what they thought about the quality of their teacher's performance.

3. That there would be anonymity with respect to participation and results of the study.

4. That teachers would receive feedback from the student responses in the form of a graph and also a tabular summation of each of the thirteen items. In addition, assurance was given that the results would be returned to the teachers personally and not shared with anyone else.

5. That there was approximately a twenty-five percent chance that those teachers participating in the initial phase would be selected to participate further in the experiment. Additionally, the teachers were told that, if selected, their role in the experiment would be further delineated at that time and continued participation would again be voluntary.

The candidates were also informed of the teacher criteria and
the setting, including the endorsement of the study by the respective boards and administrations.

The favorable response by teachers contacted for participation in the preliminary phase of the study was one hundred percent. Initially, this complete cooperation was important from a mere numbers standpoint. Had only half of the eligible teachers chosen to participate, for example, the probability of getting enough teachers capable of showing improvement on the perception scores would have been drastically reduced. Secondly, had several teachers not decided to participate it could have been reasoned that these teachers were representative of a given type of teacher. Specifically, it was feared that those teachers who perceived themselves poorly might decline participation in the experiment, thus eliminating a desirable calibre of teacher for this study. This, then, could have rendered invalid subsequent generalizations. The one hundred percent cooperation insured that certain types of teachers did not withdraw from the experiment.

Selection of the Sample

Those teachers in the three districts who were currently in their first, second, or third year of teaching and were presently assigned to a fourth, fifth, or sixth-grade self-contained classroom were asked to participate in the initial phase of the study. In an effort to control the predictable extraneous error variance the teachers selected were as homogeneous as possible.\(^1\) Teachers in their first

three years of teaching share a commonality in that three years represents the maximum time allowable as a probationary teacher (i.e. without achieving tenure status). This experience level was also consistent with the classification of experienced teachers in recent studies by Tobin\(^1\) and by Oliver and Tuckman.\(^2\) Researchers in both of these studies operationally defined experienced teachers as those having in excess of three years teaching experience. It was also assumed that teachers who had not taught extensively would be better able to adapt to the manipulation of the independent variable and hence maximize the desired experimental variance.

The additional degree of homogeneity afforded by selecting only upper elementary or intermediate school teachers (fourth, fifth, and sixth grades) was similarly considered. Discussions with professional educators, both in the field and at the university, prompted the selection of this particular level of teacher as opposed to the junior or senior high school level. Early elementary teachers, of course, were eliminated for obvious reasons inherent in their students' inability to read or sufficiently comprehend items on the data gathering device. Among the more important reasons behind the decision for upper elementary classrooms as opposed to junior or senior high school classrooms were the following:

\(^1\)op. cit., p. 8.

\(^2\)Tuckman, E. W. and Oliver, W. F., "Effectiveness of Feedback to Teachers as a Function of Source." *Journal of Educational Psychology*, LIX (August 1968), 298.
1. Elementary teachers generally experience more opportunity to plan with other professional educators and hence would be more accustomed to the type of planning required in the experiment.

2. Elementary teachers spend more time with the same group of students than their secondary counterparts and hence would have more opportunity to manipulate the independent variable.

3. Elementary teachers as a group would in all probability be less threatened and more open to methodological suggestions than the secondary school subject matter specialist.

4. Elementary teachers represent a larger population teaching similar content than in the secondary situations.

5. Elementary teachers are located in many more buildings hence the risk of contamination among the various group treatments could be reduced.

Finally, the distinction was made between self-contained and other types of classrooms. In that the assorted team teaching and cluster approaches represented a significant departure from the one teacher— one classroom concept, the preceding distinction was deemed necessary. Allowances were made, in accord with the definition of terms, for the inclusion of teachers who basically adhered to the self-contained concept, but made provisions for the teaching of special subjects by another teacher. Fifty-four teachers in the three districts met the above criteria and subsequently were asked to participate in the experiment.

In that the prime independent variable was the use of a variety of instructional methodologies (item nine, "variety in teaching," on
the SOQ) the fifty-four consenting teachers were first ranked from highest to lowest on student perception of their "variety in teaching." Beginning with the fifty-fourth ranked teacher and working upward fifteen teachers were selected who scored below the population mean on both item nine and the overall questionnaire score. A minor exception to this selection order format was made to insure that no more than two teachers from any one building were included in the experimental phase. The names of these fifteen teachers were then placed in a hat and randomly assigned by lottery to the five groups. The assignment of teachers in the same building to different groups was not allowed. Randomization in the selection of the fifteen teachers, as an aid in minimizing the unpredictable error variance, was not advisable because of the relatively few teachers receiving low student perception scores and thus the inherent potentiality of losing some of the most desirable teachers for the experiment.\(^1\)

The predicted error variance was controlled for by choosing homogeneous subjects in the original population, employing random assignment of teachers to treatments, and providing for the use of pre-posttest change scores for participating teachers as a part of the statistical design.\(^2\)

**Administration of the Instrument**

The adapted *Student Opinion Questionnaire* was administered

\(^1\)Kerlinger, *Foundations of Behavioral Research*, op. cit., p. 287.

\(^2\)loc. cit., p. 284.
personally by the researcher to the classrooms of the fifty-four participating teachers on Monday, Tuesday, and Wednesday of two consecutive weeks in February and March. This time of year more than met the suggestion of the Feedback Center that students have a minimum of nine weeks to become adequately acquainted with the teacher before completing the SOQ. Teachers were asked to leave the classroom and to complete a color-coded copy of the same questionnaire that was given to students. The investigator then administered the questionnaire to the class by reading the instructions and the questions aloud as students followed along. A slight pause was made at the conclusion of each of the thirteen items to allow students the necessary time to select an answer. This type of administration was considered advisable for this age group by all individuals involved in the preliminary planning of the study and especially the classroom teachers. This, then, warranted its administration by one individual who would read the instructions and questionnaire items in a consistent manner being careful to eliminate irregular voice emphases and fluctuations.

At the conclusion of the administration students were asked to record the last four digits of their phone number in the lower right hand corner of the back side of the questionnaire. Those without telephone numbers were asked to record the street number of their place of residence. Students were told that this would in no way

---

1Bryan, Some Observations Concerning Written Student Reactions to High School Teachers, op. cit., p. 13.
identify them as individuals, but would merely allow the researcher
to match present papers with future papers should the questionnaire
be administered to them again. This was a necessary step in the event
that the statistical treatment required an analysis of the pre-
posttest scores of individuals. Total time for the administration
of the data gathering instrument to a class was approximately fifteen
minutes.

Following the experimental treatment the same questionnaire was
readministered to the classes of the fifteen teachers who had been
randomly assigned to the various experimental treatments. This post-
test administration during the first week in May paralleled identically
procedures employed in the administration of the pretest.

Description of the Treatments

Teachers who were randomly assigned to each of the five groups
received different treatments. There were three distinct types of
treatment: The feedback (F), the log maintenance (L), and the
instructional planning (P). The combination group, utilizing all
of the above treatments (FLP) and the control group (C), together
with the treatment groups mentioned above, comprised the five different
groups.

Were all of the possible combinations of treatments to have been
isolated it would have necessitated eight different groups. To facil-
itate a visualization of these possible combinations they are here
presented using letter abbreviations for the various treatments: C, F, L, P, FL, FP, LP, FLP.
Three basic reasons prevented the inclusion of all eight groups and necessitated a priority choosing. First, to deplete the possible combinations would have required a sample of twenty-four teachers, almost half of the fifty-four teachers in the population. The probability of finding this number who would, on the pretest, score below the population means on "variety in teaching" and "overall" was not promising. Hence to include twenty-four teachers in the experimental phase would run the risk of violating the premise that the teachers in the sample were not highly perceived on either of the two above-mentioned scores. Secondly, the resources available for the experiment in time, personnel, and finances, would not allow for the planning sessions necessitated by half of the eight groups (P, FP, LP, and FLP). Thirdly, the maintenance of the checklist log was envisioned from the inception of the study as an important and inherent aid in the instructional planning sessions and consequently would have only reluctantly been deleted from that treatment. In consultation with the doctoral committee the following groups were selected as the most desirable from the standpoint of isolating variables, the most feasible from the standpoint of available resources, and the most consonant with respect to the purposes of the experiment. Selected were the control group; the feedback group; the log group; the planning and log group; and the planning, log, and feedback group. A detailed description of each of these groups follows:

Control group: The control group teachers were randomly assigned to the treatment. They were informed personally that they had been
selected to participate further in the experiment. All agreed to participate. The role which they were to play was to give the researcher permission to readminister the Student Opinion Questionnaire to their class during the first week in May.

Some consideration was given to the administration of a placebo to equalize the Hawthorne effect among the basic groups. The placebo considered was for the researcher to meet with the control group teachers individually and discuss topics foreign to the experiment. However, several reasons prompted the discarding of this consideration. First, the teachers in this group were told that they were a part of the experiment, had seen and in fact completed the questionnaire form previously, and were told that their classes would be given the questionnaire again. The situation, then, provided them first with the feeling of participation and additionally the criterion of measurement for which they could attempt improvement if they so desired. Secondly, there were two other groups who also were not getting the "personal meetings" and for whom such meetings were not desirable in conjunction with the intended purpose of the feedback and log. Thirdly, if the personal meetings proved to be part of the reason behind a change in student perception, the result could be indicative of desired supervisory behavior by principals, supervisors, and colleagues. Finally, the Hawthorne effect among those responsible for scoring, namely the students, was equalized among groups because teacher involvement in an experimental study was not made known to students.
**Feedback group:** The feedback group of teachers was informed personally that they had been selected to participate further in the experiment and that they had been randomly assigned to this treatment. All agreed to participate further. These teachers were given the feedback information (Appendixes B, C, and D) and told that their classes would be readministered the SOQ during the first week in May.

**Log group:** Teachers in the log group were informed personally that they had been selected to participate further in the experiment and that they had been randomly assigned to this treatment. All agreed to participate further. The directions for operationalizing the checklist log were explained individually in detail to each of these teachers. Each teacher was given the necessary forms to maintain the log for the first four weeks of the experiment and also the phone number of the investigator should problems have arisen. One week into the experiment a check was made with each of the teachers to uncover any procedural difficulties. Midway through the experiment another check was made and the necessary forms delivered to complete the final four weeks of the experiment. No procedural difficulties were experienced by any of these teachers during any part of the experiment.

**Planning group:** Teachers in the planning group were also informed personally of their selection and random assignment to this particular treatment group. All agreed to participate further in the experiment. It was explained that teachers in this group were to maintain a checklist log of their instructional methodologies and also to meet individually with a person skilled in the employment
of educational methodologies for approximately one hour weekly for eight weeks. The directions for operationalizing the checklist log were explained in a fashion similar to that of the log group. The planning sessions began the week prior to March 16, which marked the beginning of the experiment. During the hour of weekly planning the teacher and the consultant worked together planning instructional methodology and teaching strategy for the ensuing week. Paramount attention was given to the use of a variety of educational methods, both from day to day and also during any given day. These sessions did not try to pit one method against another or attempt to show the efficacy of a single method of teaching. Rather, the planning that took place focused on teachers using many educational methods at their disposal, both human and technological, to bear on the situation at hand. Obviously, such items as the availability of materials, the time allotted for presentation, the collective characteristics of the students and the teacher, and the subject matter to be presented played significant roles in the selection of the instructional methods and the teacher's strategy. Also of no small consequence was the necessary rapport established between the teacher and the consultant. Understandably, the researcher had no authority with which to force the use of various methodologies. His role was most frequently to suggest the available methods for presentation and whenever possible to encourage the teacher to try different and varied methodologies. The analysis of the checklist log was invaluable in this regard. On occasion the consultant did assist the
teacher in finding and preparing materials and in the operation of equipment prior to its use in the classroom. On no occasion did the consultant take part in any classroom presentation. The opportunity to plan for instruction in this way was, in the opinion of the consultant, a mutually satisfying and successful endeavor.

**Combination group:** All teachers randomly assigned to the combination group agreed to participate further in the experiment. These teachers were informed personally of their selection and their further role in the experiment was outlined. Teachers in this group received the combined treatments of groups A, B, and C. They were given the feedback from student pretests, were instructed in the use of the checklist log, and maintained it during the experiment. Finally, these teachers met with a person skilled in the employment of educational methodologies for one hour weekly during the experiment. These sessions were similar to those in group B except that it was also possible here to incorporate the feedback information in the planning. Again, the researcher believed mutually rewarding the challenge of meaningfully planning for the process of instruction.

**Collection and Organization of the Data**

After the administration of the pretest a graphical teacher profile (Appendix B) and a tabular summation (Appendix C) of student responses were compiled for each of the fifty-four participating teachers. In accord with the design of the experiment the three teachers in group A, the feedback group, and the three teachers in
group D, the combination group, were given feedback information as part of their prescribed treatments. The pretest information for the other forty-eight teachers was withheld until the conclusion of the experiment. Following the administration of the posttest to the classes of the fifteen teachers participating in the experimental phase of the study, a graphical teacher profile and a tabular summation of student posttest scores were compiled for these experimental phase teachers. Teachers participating in this experimental phase were then given both pre- and posttest feedback information. The thirty-nine teachers participating only in the initial phase of the study were also given at this time the pretest information previously compiled for them. The feedback information here described was promised to all participants when they agreed to participate in the study.

The pre-and posttest data from the fifteen teachers participating in the experimental phase of the study were organized for analysis. Students' pre- and posttest responses were matched by using the telephone code number the students had provided. The responses of those students who were absent from either the pre- or posttest administration were discarded leaving only the responses of students who were present for both the pre- and posttest administrations. Data from the students' pre- and posttest responses were keypunched into 370 data processing cards. A computer program was selected that would provide the statistical information necessary
for analysis according to the research design described earlier in this chapter (Appendix G). Chapter IV will include the presentation and analyses of these data.
CHAPTER IV

PRESENTATION AND ANALYSIS OF THE DATA

This chapter will include the presentation and analysis of the data which were collected to carry out the objectives of this experiment. The primary objective was to determine if teachers could modify student perceptions of their teaching performances by employing multiple educational methodologies in classroom presentations. The statistical model used to test for significant differences from pre- to posttest measures was the t-test.

A secondary objective of the experiment was to ascertain the relative effectiveness of four experimental treatments in helping a teacher to modify student perceptions of his teaching performance. To test for differences among groups a one-way analysis of variance statistical model was used. This model simultaneously compares the means of two or more groups by contrasting the variance between groups with the variance within groups. Results from an analysis of variance are presented in the form of F-ratios.

The format of this chapter follows the sequence of the "Major Questions to be Investigated" and the "Specific Questions to be Investigated" presented in Chapter III. Findings for questions 1a-1e and 3a-3e are reported in the form of t-tests. Analysis of variance models are used to report the findings for questions 2a-2j and 4a-4j.

1Kerlinger, Foundations of Behavioral Research, op. cit., p. 192.
Students' Overall Perceptions of Teachers

Major questions one and two of this study dealt with students' overall perceptions of teachers as elicited by the thirteen criterion items of the adapted Student Opinion Questionnaire. Prior to presenting the data which could provide answers to these questions it was advisable to determine if there were significant differences between groups on the overall pretest measure. This determination was needed before it could be assumed that the random assignment of teachers to groups did in fact provide starting points which were not significantly different. The data in Table 3 indicate that the overall mean scores for the five groups ranged from a low of 3.37 to a high of 3.63. Clearly shown in the F-ratio of Table 4, however, is the fact that this difference is not a significant one. The F-ratio of 1.80 for the analysis of variance does not exceed the table value of 2.40 for the distribution of F at the .05 level of significance. Thus it was possible to assume that the random assignment of teachers to groups did in fact provide starting points which were not significantly different.

Major question one with its five component parts sought to determine whether there was a significant difference in overall student perception of teacher performance from pre- to posttest measure for each of the five groups.

Question 1a: Was the overall student perception of teacher performance for teachers in group A (written feedback) more favorable on the posttest measure than on the pretest measure?
Table 3
Summary Data for Overall Pretest Measure

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>77</td>
<td>3.54</td>
<td>.588</td>
</tr>
<tr>
<td>B</td>
<td>77</td>
<td>3.63</td>
<td>.570</td>
</tr>
<tr>
<td>C</td>
<td>72</td>
<td>3.45</td>
<td>.574</td>
</tr>
<tr>
<td>D</td>
<td>67</td>
<td>3.37</td>
<td>.738</td>
</tr>
<tr>
<td>E</td>
<td>77</td>
<td>3.53</td>
<td>.707</td>
</tr>
</tbody>
</table>

Table 4
Analysis of Variance for Overall Pretest Measure

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups variance</td>
<td>4</td>
<td>501.628</td>
<td>125.407</td>
<td>1.800*</td>
</tr>
<tr>
<td>Within groups variance</td>
<td>365</td>
<td>25,424.382</td>
<td>69.655</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td>369</td>
<td>25,926.011</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table value is 2.40 at the .05 level of significance
The data suggest that teachers in the feedback group were not more favorably perceived by their students after the experimental treatment of receiving student feedback regarding their teaching performances. On the contrary, the results shown in Table 5 approach a significant difference, but in the direction opposite to that expected. The overall mean score for teachers in this group decreased from an initial mean of 3.54 to a final mean of 3.37.

Table 5

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Pretest mean</th>
<th>Posttest mean</th>
<th>t-score of change</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>77</td>
<td>3.54</td>
<td>3.37</td>
<td>1.97</td>
<td>N.S.a</td>
</tr>
<tr>
<td>B</td>
<td>77</td>
<td>3.63</td>
<td>3.68</td>
<td>0.67</td>
<td>N.S.</td>
</tr>
<tr>
<td>C</td>
<td>72</td>
<td>3.45</td>
<td>3.56</td>
<td>1.36</td>
<td>N.S.</td>
</tr>
<tr>
<td>D</td>
<td>67</td>
<td>3.37</td>
<td>3.35</td>
<td>0.13</td>
<td>N.S.</td>
</tr>
<tr>
<td>E</td>
<td>77</td>
<td>3.53</td>
<td>3.47</td>
<td>0.93</td>
<td>N.S.</td>
</tr>
</tbody>
</table>

*aTable value is 2.00 at the .05 level of significance*

Question 1b: Was the overall student perception of teacher performance for teachers in group B (log maintenance) more favorable on the posttest measure than on the pretest measure?

The data shown in Table 5 indicate that student perceptions were not significantly more favorable on the overall posttest measure. The experimental treatment of log maintenance for this group of teachers...
was not significantly successful in helping them improve student perceptions of their overall teaching performances. The increase in overall mean score from 3.63 to 3.68 produced a $t$-score of 0.67, short of the 2.00 value needed for significance at the .05 level.

Question 1c: Was the overall student perception of teachers in group C (employment of multiple educational methodologies and log maintenance) more favorable on the posttest measure than on the pretest measure?

Table 5 indicates that the overall mean scores for teachers in group C increased from 3.45 to 3.56. This was the best gain made by any of the four experimental groups. The $t$-score of 1.36, however, did not exceed the value necessary for significance at the .05 level. According to the data student perceptions of teachers employing multiple educational methodologies and log maintenance were not significantly more favorable after the experimental treatment.

Question 1d: Was the overall student perception of teacher performance for teachers in group D (employment of multiple educational methodologies, log maintenance, and written feedback) more favorable on the posttest measure than on the pretest measure?

The data in Table 5 indicate that the overall mean scores for teachers in this group decreased from 3.37 to 3.35. Obviously, this slight decrease was not a significant difference. Clearly dictated by the results is a negative response to the above question.

Question 1e: Was the overall student perception of teacher performance for teachers in group E (control) similar on the pre- and posttest measures?
Reference to Table 5 indicates that there was a similarity on the pre- and posttest measures for teachers in the control group. The data show that a decrease in overall mean score from 3.53 to 3.47 occurred during the course of the experiment. This decrease in mean score, however, was not statistically significant.

A more meaningful picture of differences between the various group pre- and posttest measures could perhaps be realized by computing $t$-tests of change scores (pre- to posttest) based on differences in the control group rather than the departure from zero as done in Table 5. The results of such computations are shown in Table 6. Since the control group change in mean was a negative one the $t$-scores of those groups showing a positive change (groups B and C) might be expected to be larger. Correspondingly, the $t$-scores for those groups showing a negative change (groups A and D) might be expected to be smaller than the $t$-scores shown in Table 5 which did not take into account the differences in the control group. When analyzed in the manner presented in Table 6 the results still indicate no significant results at the .05 level.

Major question two, with its ten component parts, sought to determine whether there was a significant difference between groups on the overall posttest measure.

The data in Tables 7 and 8 indicate that the analysis of variance computed for the overall posttest measure revealed a significant difference between group means. A further analysis of this difference

---

1loc. cit., p. 309.
Table 6

*t*-tests of Change Scores (Pre- to Posttest) for Overall Means
Based on Differences in the Control Group

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Pretest Mean</th>
<th>Posttest Mean</th>
<th>t-score of Change</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>77</td>
<td>3.54</td>
<td>3.37</td>
<td>0.92</td>
<td>N.S.</td>
</tr>
<tr>
<td>B</td>
<td>77</td>
<td>3.63</td>
<td>3.68</td>
<td>1.13</td>
<td>N.S.</td>
</tr>
<tr>
<td>C</td>
<td>72</td>
<td>3.45</td>
<td>3.56</td>
<td>1.63</td>
<td>N.S.</td>
</tr>
<tr>
<td>D</td>
<td>67</td>
<td>3.37</td>
<td>3.35</td>
<td>0.54</td>
<td>N.S.</td>
</tr>
<tr>
<td>E</td>
<td>77</td>
<td>3.53</td>
<td>3.47</td>
<td>0.00</td>
<td>N.S.</td>
</tr>
</tbody>
</table>

Table 7

Summary Data for Overall Posttest Measure

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Posttest Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>77</td>
<td>3.37</td>
<td>.798</td>
</tr>
<tr>
<td>B</td>
<td>77</td>
<td>3.68</td>
<td>.653</td>
</tr>
<tr>
<td>C</td>
<td>72</td>
<td>3.56</td>
<td>.690</td>
</tr>
<tr>
<td>D</td>
<td>67</td>
<td>3.35</td>
<td>.672</td>
</tr>
<tr>
<td>E</td>
<td>77</td>
<td>3.47</td>
<td>.802</td>
</tr>
</tbody>
</table>
Table 8

Analysis of Variance for Overall Posttest Measure

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups variance</td>
<td>4</td>
<td>936.742</td>
<td>234.185</td>
<td>2.58a</td>
</tr>
<tr>
<td>Within groups variance</td>
<td>365</td>
<td>33,096.296</td>
<td>90.674</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td>369</td>
<td>34,033.039</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*aTable value is 2.40 at the .05 level of significance

was made possible by computing the overall posttest between group
\(t\)-values. These data, shown in Table 9, indicate that the significant
\(F\)-ratio computed for Table 8 can be attributed in large part to the
differences between the posttest means of groups A and B and groups
B and D. However, a reanalysis of the data in Tables 3 and 4

Table 9

Overall Posttest Between Group \(t\)-values

<table>
<thead>
<tr>
<th>Group</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>2.636a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1.499</td>
<td>-1.144</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>-0.135</td>
<td>-2.950a</td>
<td>-1.733</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>0.731</td>
<td>-1.827</td>
<td>-0.728</td>
<td>0.892</td>
<td>0.000</td>
</tr>
</tbody>
</table>

*aTable value is 2.61 at the .01 level of significance

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
indicates that the overall pretest mean scores were not identical. Consequently, an analysis of variance using pre- to posttest change scores rather than an analysis of variance using posttest scores alone might give a more realistic indication of changes accruing from the experimental treatment. The design of the experiment whereby the pre- and posttest scores of individuals were paired made possible such an analysis using change score data.

These data are presented in Table 10 and indicate that there were no significant differences in the effectiveness of one group over another in changing student perceptions of teacher performance.

Table 10

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups variance</td>
<td>4</td>
<td>590.976</td>
<td>174.744</td>
<td>1.923*</td>
</tr>
<tr>
<td>Within groups variance</td>
<td>365</td>
<td>28,048.457</td>
<td>76.845</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td>369</td>
<td>28,639.433</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

aTable value is 2.39 at the .05 level of significance

Based on the lack of significance of the F-ratio shown in Table 10, responses to questions 2a-2j, comparing each treatment group in turn with every other treatment group would all be similar: No
treatment was found to be significantly more effective than any other in changing overall student perceptions of teacher performance.

Students' Perceptions of Their Teachers' "Variety in Teaching"

Major questions three and four of this study dealt with student perception of teacher performance relating specifically to the teachers' "variety in teaching" (item nine, SOQ). Prior to presenting the data which could provide answers to these questions it was deemed advisable to determine if there were significant differences between groups on SOQ item nine, "variety in teaching" of the pretest measure. This determination was needed before it could be assumed that the random assignment of teachers to groups did in fact provide starting points which were not significantly different.

The data in Table 11 indicate that the mean scores on SOQ item nine, "variety in teaching," for the five groups ranged from a low of 2.97 to a high of 3.54. Moreover, the F-ratio of 2.61 shown in Table 12 exceeds slightly the table value of 2.40 for the distribution of F at the .05 level of significance. These data suggest that there was a significant difference in group means for SOQ item nine on the pretest measure. Thus it was impossible to assume that the random assignment of teachers to groups did in fact provide starting points which were not significantly different.

A further analysis of this difference is made possible by computing the between group t-values for SOQ item nine, "variety in teaching." These data, shown in Table 13, indicate that the significant F-ratio computed for Table 12, can be attributed in large
part to the differences between the pretest means of groups B and D and groups D and E. This fact will be taken into consideration when analyzing group posttest means on SOQ item nine, "variety in teaching."

Table 11
Summary Data for SOQ Item Nine, "Variety in Teaching," of the Pretest Measure

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>77</td>
<td>3.22</td>
<td>1.42</td>
</tr>
<tr>
<td>B</td>
<td>77</td>
<td>3.54</td>
<td>1.17</td>
</tr>
<tr>
<td>C</td>
<td>72</td>
<td>3.17</td>
<td>1.19</td>
</tr>
<tr>
<td>D</td>
<td>67</td>
<td>2.97</td>
<td>1.52</td>
</tr>
<tr>
<td>E</td>
<td>77</td>
<td>3.54</td>
<td>1.26</td>
</tr>
</tbody>
</table>

Table 12
Analysis of Variance for SOQ Item Nine, "Variety in Teaching," of the Pretest Measure

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups variance</td>
<td>4</td>
<td>18.331</td>
<td>4.582</td>
<td>2.61a</td>
</tr>
<tr>
<td>Within groups variance</td>
<td>365</td>
<td>639.368</td>
<td>1.751</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td>369</td>
<td>657.700</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

αTable value is 2.40 at the .05 level of significance
Major question three, with its five component parts, sought to determine whether there was a significant difference in student perceptions of the teachers' "variety in teaching" from pre- to posttest measure for each of the five groups.

Table 13

Pretest Between Group t-values for SOQ Item Nine, "Variety in Teaching"

<table>
<thead>
<tr>
<th>Group</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>1.539</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>-0.249</td>
<td>-1.946</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>-1.016</td>
<td>-2.549*</td>
<td>-0.847</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>1.488</td>
<td>0.000</td>
<td>1.867</td>
<td>2.465*</td>
<td>0.000</td>
</tr>
</tbody>
</table>

*Table value is 1.98 at the .05 level of significance

Question 3a: Was student perception of teacher performance regarding "variety in teaching" for teachers in group A (written feedback) more favorable on the posttest than on the pretest measure?

The data suggest that teachers in the feedback group were not more favorably perceived by their students after the experimental treatment of receiving student feedback regarding their teaching performance. On the contrary the results shown in Table 14 indicate a drop in mean score for SOQ item nine, "variety in teaching," from 3.22 to 3.14. This was not, however, a statistically significant change.
Question 3b: Was student perception of teacher performance regarding "variety in teaching" for teachers in group B (log maintenance) more favorable on the posttest than on the pretest measure?

Table 14

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Pretest Mean</th>
<th>Posttest Mean</th>
<th>t-score of Change</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>77</td>
<td>3.22</td>
<td>3.14</td>
<td>0.43</td>
<td>N.S.</td>
</tr>
<tr>
<td>B</td>
<td>77</td>
<td>3.54</td>
<td>4.54</td>
<td>6.71</td>
<td>.001a</td>
</tr>
<tr>
<td>C</td>
<td>72</td>
<td>3.17</td>
<td>4.25</td>
<td>7.09</td>
<td>.001a</td>
</tr>
<tr>
<td>D</td>
<td>67</td>
<td>2.97</td>
<td>4.07</td>
<td>6.70</td>
<td>.001a</td>
</tr>
<tr>
<td>E</td>
<td>77</td>
<td>3.54</td>
<td>3.36</td>
<td>0.98</td>
<td>N.S.</td>
</tr>
</tbody>
</table>

*Table value is 3.46 at the .001 level of significance

According to the data in Table 14 teachers in the log group were more favorably perceived by their students on the posttest measure than on the pretest measure. The change in mean score on SQQ item nine from 3.54 to 4.54 produced a t-score of 6.71 which exceeded the table value of 3.46 necessary for significance at the .001 level. These data indicate that teachers maintaining the checklist log effected a significant improvement in student perception of their "variety in teaching" during the course of the experiment.

Question 3c: Was student perception of teacher performance regarding "variety in teaching" for teachers in group C (employment
of educational methodologies and log maintenance) more favorable on the posttest measure than on the pretest measure?

The data in Table 14 reveal that student perception was indeed more favorable on the posttest measure than on the pretest measure. The improvement of the mean score on SOQ item nine from 3.17 to 4.25 produced a t-score of 7.09 which was significant at the .001 level. Thus, teachers who employed a variety of educational methodologies and maintained the checklist log showed significantly improved student perceptions of their "variety in teaching."

Question 3d: Was student perception of teacher performance regarding "variety in teaching" for teachers in group D (employment of multiple educational methodologies, log maintenance, and written feedback) more favorable on the posttest measure than on the pretest measure?

According to the data in Table 14 student perception for teachers in group D was more favorable on the posttest than on the pretest. The pretest mean of 2.97 on SOQ item nine was increased to 4.07 on the posttest. Significant at the .001 level was the resultant t-score of 6.70. The data suggest that teachers in the combination group were able to significantly change student perceptions of their "variety in teaching."

Question 3e: Was student perception of teacher performance regarding "variety in teaching" for teachers in group E (control) similar on the pre- and posttest measures?

The data in Table 14 indicate that the mean score for teachers in the control group decreased from 3.54 to 3.36. However, the 0.98
\( t \)-score computed for this decrease did not exceed the table value of 2.00 needed for significance at the .05 level. Thus, an affirmative answer can be given for question 3e.

A more meaningful picture of differences between the various group pre- and posttest measures regarding SOQ item nine, "variety in teaching," could perhaps be realized by computing \( t \)-tests of change scores (pre- to posttest) based on differences in the control group rather than the departure from zero as done in Table 14. The results of such computations are displayed in Table 15. When analyzed in this manner the results still show significance at the .001 level for groups B, C, and D.

Table 15
\( t \)-tests of Change Scores (Pre- to Posttest) for SOQ Item Nine, "Variety in Teaching," Based on Differences in the Control Group

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Pretest Mean</th>
<th>Posttest Mean</th>
<th>( t )-score of Change</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>77</td>
<td>3.22</td>
<td>3.14</td>
<td>0.40</td>
<td>N.S.</td>
</tr>
<tr>
<td>B</td>
<td>77</td>
<td>3.54</td>
<td>4.54</td>
<td>4.95</td>
<td>.001a</td>
</tr>
<tr>
<td>C</td>
<td>72</td>
<td>3.17</td>
<td>4.25</td>
<td>5.21</td>
<td>.001a</td>
</tr>
<tr>
<td>D</td>
<td>67</td>
<td>2.97</td>
<td>4.07</td>
<td>5.10</td>
<td>.001a</td>
</tr>
<tr>
<td>E</td>
<td>77</td>
<td>3.54</td>
<td>3.36</td>
<td>0.00</td>
<td>N.S.</td>
</tr>
</tbody>
</table>

\(^a\)Table value is 3.46 at the .001 level of significance

\(^1\)ibid.
Major question four, with its ten component parts, sought to determine whether there was a significant difference between groups on the posttest measure for SOQ item nine, "variety in teaching."

The data in Tables 16 and 17 indicate that the analysis of variance computed for the SOQ item nine posttest measure revealed a significant difference between groups.

Table 16
Summary Data for SOQ Item Nine, "Variety in Teaching," of the Posttest Measure

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Posttest Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>77</td>
<td>3.14</td>
<td>1.42</td>
</tr>
<tr>
<td>B</td>
<td>77</td>
<td>4.54</td>
<td>0.64</td>
</tr>
<tr>
<td>C</td>
<td>72</td>
<td>4.25</td>
<td>1.06</td>
</tr>
<tr>
<td>D</td>
<td>67</td>
<td>4.07</td>
<td>1.04</td>
</tr>
<tr>
<td>E</td>
<td>77</td>
<td>3.36</td>
<td>1.39</td>
</tr>
</tbody>
</table>

A further analysis of the difference is made possible by computing the posttest between group t-values for SOQ item nine, "variety in teaching." These data, shown in Table 18, indicate that the significant F-ratio computed for Table 17 can be attributed to the differences between the posttest means of groups A and B, A and C, A and D, B and C, B and D, B and E, C and E, and D and E.

However, a reanalysis of the data in Table 12 indicates that there were significant differences on the analysis of variance of the
### Table 17

Analysis of Variance for SOQ Item Nine, "Variety in Teaching," of the Posttest Measure

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups variance</td>
<td>4</td>
<td>108.778</td>
<td>27.194</td>
<td>20.32$^a$</td>
</tr>
<tr>
<td>Within groups variance</td>
<td>365</td>
<td>488.464</td>
<td>1.338</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td>369</td>
<td>597.243</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$Table value is 3.37 at the .01 level of significance

### Table 18

Posttest Between Group $t$-values for SOQ Item Nine, "Variety in Teaching"

<table>
<thead>
<tr>
<th>Group</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>7.856$^b$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>5.319$^b$</td>
<td>-2.059$^a$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>4.400$^b$</td>
<td>-3.297$^b$</td>
<td>-0.974</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>0.969</td>
<td>-6.759$^b$</td>
<td>-4.328$^b$</td>
<td>3.415$^b$</td>
<td>0.000</td>
</tr>
</tbody>
</table>

$^a$Table value is 1.98 at the .05 level of significance

$^b$Table value is 2.61 at the .01 level of significance
pretest measure for SOQ item nine, "variety in teaching." Consequently, an analysis of variance using pre- to posttest change scores rather than an analysis of variance using posttest scores alone could give a more realistic indication of changes accruing from the experimental treatment. The design of the experiment whereby the pre- and posttest scores of individuals were paired made possible such an analysis using change score data.

These data are presented in Table 19 and indicate that there were significant differences in the effectiveness of one group treatment over another in changing student perceptions of teacher performance. The analysis of variance produces an F-ratio of 15.34, exceeding the table value of 3.37 needed for significance at the .01 level.

Table 19

Analysis of Variance for SOQ Item Nine, "Variety in Teaching," of the Change Score Measure

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups variance</td>
<td>4</td>
<td>128.187</td>
<td>32.046</td>
<td>15.335*</td>
</tr>
<tr>
<td>Within groups variance</td>
<td>365</td>
<td>762.755</td>
<td>2.089</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td>369</td>
<td>890.943</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table value is 3.37 for significance at the .01 level.
The results of the change score between group $t$-values for "variety in teaching" shown in Table 20 confirm the above analysis and provide insight on differences between specific paired combinations of groups. With regard to specific questions 4a-4j, comparing each treatment group in turn with every other treatment group, a summarization would be appropriate.

Tables 15, 19, and 20 clearly indicate that groups B, C, and D were significantly more effective (.01 level) than groups A and E in effecting change in student perceptions of their teachers' "variety in teaching." Furthermore, no significant difference was found between groups A and E. Finally, no significant differences were found between groups B, C, and D, all of which were successful (.001 level of significance) in effecting a change from pre- to posttest measure.

A more detailed analysis of each paired comparison (specific questions 4a-4j) follows:

Table 20

<table>
<thead>
<tr>
<th>Group</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>4.612$^a$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>4.858$^a$</td>
<td>-0.390</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>4.791$^a$</td>
<td>-0.471</td>
<td>-0.094</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>0.400</td>
<td>4.950$^a$</td>
<td>5.206$^a$</td>
<td>5.099$^a$</td>
<td>0.000</td>
</tr>
</tbody>
</table>

$^a$Table value is 3.46 at the .001 level of significance
Question 4a: Was the written feedback of student perception of teacher performance (group A) more effective than no treatment at all in helping a teacher to modify student perception of his "variety in teaching"?

The results of the t-test shown in Table 20 clearly indicate that there was not a significant difference between groups A and E in modifying student perception of the teachers' "variety in teaching." An examination of the pre- and posttest means for the two groups displayed in Table 15 reveals that both groups showed a decrease in mean score with the feedback group less negative than the control group.

Question 4b: Was the maintenance of the checklist log of various educational methodologies (group B) more effective than no treatment at all in helping a teacher to modify student perception of his "variety in teaching"?

Displayed in Table 20 is a t-value of 4.950, indicating a difference in group means significant at the .001 level. While group E decreased in mean score from pre- to posttest (3.54 to 3.36), Table 15 shows that quite the contrary was true for group B. The log group increased from a pretest mean of 3.54 to a posttest mean of 4.54. An affirmative answer can be given for question 4b.

Question 4c: Was the employment of multiple educational methodologies and log maintenance (group C) more effective than no treatment at all in helping a teacher to modify student perception of his "variety in teaching"?

The t-values in Table 20 reveal a difference in group means significant at the .001 level. Mean scores in Table 15 show an increase
from 3.17 to 4.25 during the course of the experiment. The planning group treatment was indeed more effective than no treatment at all in helping a teacher to modify student perception of his "variety in teaching."

Question 4d: Was the employment of a combination of feedback, log maintenance, and multiple educational methodologies (group D) more effective than no treatment at all in helping a teacher to modify student perception of his "variety in teaching"?

Once again an affirmative answer can be given based on the significant t-value shown in Table 20. The t-value of 5.099 is significant at the .001 level. Increasing from a pretest mean of 2.97 to a posttest mean of 4.07 as shown in Table 15, the combination group was clearly more effective than no treatment at all in helping a teacher to modify student perception of his "variety in teaching."

Question 4e: Was the maintenance of a checklist log of various educational methodologies (group B) more effective than written feedback alone (group A) in helping a teacher to modify student perception of his "variety in teaching"?

The data in Table 20 indicate a significant difference between the two group means at the .001 level. While the feedback group decreased in mean score from 3.22 to 3.14 from pre- to posttest, the results in Table 15 reveal that the log group increased in mean score from 3.54 to 4.54. Consequently, an affirmative answer can be given for question 4e.

Question 4f: Was the employment of multiple educational methodologies supplemented with log maintenance (group C) more effective
than written feedback alone (group A) in helping a teacher to modify student perception of his "variety in teaching"?

A t-value of 4.885 for this comparison again exceeds the table value of 3.46 needed for significance at the .001 level. The planning group showed a large increase in mean score from pre- to posttest while the feedback group decreased slightly. The results indicate that the planning group was more effective than the feedback group in helping a teacher to modify student perception of his "variety in teaching."

Question 4g: Was the employment of multiple educational methodologies supplemented with log maintenance and written feedback of student perception of teacher performance (group D) more effective than written feedback alone (group A) in helping a teacher to modify student perception of his "variety in teaching"?

Clearly demonstrated in Table 20 is the greater effectiveness of the combination group in comparison with the feedback group in helping a teacher to modify student perception of his "variety in teaching." The t-value of 4.791 is significant at the .001 level. The large increase in mean score of the combination group (2.97-4.07) from pre- to posttest while the feedback group was decreasing slightly accounted for the significant difference.

Question 4h: Was the employment of multiple educational methodologies supplemented with log maintenance (group C) more effective than the maintenance of a checklist log alone (group B) in helping a teacher to modify student perception of his "variety in teaching"?
The results of Table 20 show that there was not a significant difference between the two groups. The data in Table 15 indicate that both groups were very effective (.001 level) in improving from pre- to posttest. The planning group and the log group were about equally effective in helping a teacher to modify student perception of his "variety in teaching."

Question 4i: Was the employment of multiple educational methodologies supplemented with log maintenance and written feedback of student perception of teacher performance (group D) more effective than the maintenance of a checklist log alone (group B) in helping a teacher to modify student perception of his "variety in teaching"?

A negative response must be given to the above question. The \( t \)-value of 0.471 shown in Table 20 is far less than that needed for significance. Although the data in Table 15 indicate that both group treatments were effective (.001) in helping a teacher to modify student perception of his "variety in teaching," neither was significantly more effective than the other.

Question 4j: Was the employment of multiple educational methodologies supplemented with log maintenance and written feedback of student perception of teacher performance (group D) more effective than the employment of multiple educational methodologies and log maintenance alone (group C) in helping a teacher to modify student perception of his "variety in teaching"?

Once again the comparison was made between two highly effective group treatments as shown in Table 15. The \( t \)-value of 0.094 in Table
indicates that there was not a significant difference between the two groups. The combination and planning groups were about equally effective in helping a teacher to modify student perception of his "variety in teaching."

In addition to the above presentation of data, student perception averages for all questionnaire items, pre and post, will be included in tabular form in Appendix H. This summation provides information by individual teacher and by the five group treatments.

Open-Ended Participant Responses

At the conclusion of the experiment each of the teachers in the experimental groups was asked to comment about the effect, if any, of the experimental "treatment" on their teaching during the course of the field experiment. An open-ended form was used for gathering this information (Appendixes I, J, K, and L). The data collected from this form were summarized for each experimental group and are here presented.

Teachers in the feedback group, group A, were first asked to comment about the effect, if any, of the feedback on their teaching during the course of the experiment. One of the teachers in this group indicated that the feedback was used first as a means of analyzing teacher performance and secondly as an aid in attempting to improve in areas where that performance was perceived poorly by students. Two of the teachers indicated that the information was used very little. One teacher in this group, whose performance was perceived much less favorably on the posttest than on the pretest,
indicated that administrative pressure (not related to the experiment) to improve classroom control became a primary objective. Students in this classroom perceived an improvement in classroom control as revealed on the posttest, but with a resultant less favorable overall perception.

Teachers in group A were also asked for general comments about the field experiment. While two of these teachers revealed that the feedback information was helpful and reflected accurately the opinions of their students, one of the teachers, who tried nothing different as a result of the feedback, felt that the student interpretation of some questions was "somewhat different than that of adults."

Teachers in group B, the log group, were asked to comment first on the effect of the checklist log on their teaching during the course of the field experiment. The responses were all similar in nature. These teachers indicated that the maintenance of the log made them aware of what methods they were using and served as a challenge to employ a greater variety of methods in their teaching. Some of the comments recorded were as follows:

"The checklist log made me aware of exactly what methods and equipment I was using."

"I found that I tried to use different types of teaching techniques because I felt I was in a rut doing the same things over and over."

"Being confronted with the checklist log each afternoon and its wide variety of teaching devices and materials made it challenging to try as many as possible."

The general comments of teachers in group B about the field experiment were varied. Two of the teachers expressed again the value of
experimenting with a variety of methods. One teacher said:

"Some methods like role-playing helped teach attitudes rather than the skill. Because of this experience I find I am teaching more individual children rather than facts of a subject."

The third teacher in this group commented on the pretest and indicated that some students reacted too strongly to the idea of "grading" the teacher.

Finally, all teachers in this group were asked whether they had employed a greater variety of instructional methodologies during than prior to the experiment. All answered affirmatively with one teacher indicating that more instructional devices would have been used had they been available or in working order.

Teachers in the planning group, group C, were first asked to comment on the effect, if any, of the checklist log on their teaching. The responses were similar to those expressed by teachers in group B in that the awareness of what was being used and the challenge to employ a greater variety of methods were both mentioned. One of the teachers offered this comment:

"Over a period of time you would notice a trend in methods and equipment that was employed in teaching. The checklist helped you to notice these trends and to see where improvement or change was needed."

In response to the effect of the weekly planning sessions, teachers in this group indicated that the exploration of new ideas and methods for lesson presentations was most beneficial. Some of the comments follow:

"The weekly planning sessions played an important part in developing new ideas and methods to be used, which otherwise would have been skipped."
"Some good ideas were suggested in the planning sessions. They helped promote more variety."

Only one of the teachers in this group responded in a general way about the experiment. Recorded was this reaction:

"I feel that the experiment was a huge success. In forcing myself to use a greater number of teaching methods and materials than would normally have been used, I found that every day I was trying to do something different in class. This not only increased my interest in teaching, but I also feel increased the interests of my students."

All teachers in group C responded affirmatively to the question, "Did you, during the course of the experiment, employ a greater variety of instructional methodologies than prior to the experiment"?

Teachers in group D, the combination group, were asked to comment about the effect of the feedback, the checklist log, and the weekly planning sessions on their teaching during the course of the experiment. The reactions of these teachers to the feedback information were all favorable. Primarily its benefit was viewed as providing an insight into areas on which efforts for improvement could be concentrated. Also mentioned was the self-evaluation which followed the inspection of the feedback information.

Similar to the comments of teachers in group B were the remarks of group D teachers about the checklist log. The awareness the log created of what methods the teachers were using and also of what methods were available was the most popular response. There were some additional reflections:

"I believe using the checklist log has helped me provide a more well-rounded curriculum and more learning experiences for my children."
"With the day to day record I found myself doing a better job of planning and enjoying that task for the first time. The result was a greater variety of teaching techniques."

"...when the experiment was over I missed the picture of my efforts which the checklist log provided."

Several comments were made about the weekly planning sessions by teachers in group D. In addition to the opportunity these sessions afforded for exploring the various possibilities for lesson presentations teachers indicated that these sessions were helpful in planning for the implementation of some of these methodologies in their classes. It was also noted that these planning sessions aided in a familiarization of methods and materials not previously used by the teacher. Some of the more pertinent reactions were recorded as follows:

"The weekly planning sessions were beneficial in that I received helpful suggestions in areas where I was not certain how to proceed. I believe the sessions helped familiarize me with methods I had not previously used. I used the opaque projector for the first time and attempted several science experiments which I had not previously done."

"These planning sessions proved to be of value in that some suggestions were offered as to how I could implement some of the teaching methods and instructional devices in the log."

The general comments about the experiment by teachers in this group were very favorable. Perhaps the best indication of these teachers' reactions are the direct quotations of their comments about the experiment:

"In all, I would say the experiment has been a most beneficial experience for me and, hopefully, for the children. I feel the children gained more incentive for learning and more ability to think and find out things for themselves."
"I believe this was a very worthwhile experience for me as a teacher. I believe it has had a great effect on my plans in the future as to how I can make the classroom more interesting for my students."

"I was very glad to have participated in the experiment. The experiment forced me to evaluate what I was doing in the classroom."

Finally, the teachers in this combination group were asked whether they had employed a greater variety of instructional methodologies during than prior to the experiment. All responded in the affirmative.

In addition to the above presentation of data teachers' self-perception averages for all questionnaire items, pre and post, will be included in tabular form in Appendix M. This summation provides information by individual teacher and by the five group treatments.

The Checklist Log

In addition to serving as a vehicle for motivating teachers to employ a variety of instructional methodologies the checklist log also indicated the degree to which teachers used the various educational methodologies. Nine teachers, three each in groups B, C, and D, maintained the log as a part or all of their experimental treatment. To give some indication of past performance with respect to the instructional methodologies these nine teachers were asked at the beginning of the experiment to recall and record the methodologies used during the week prior to the commencement of the experiment. During the next eight weeks the teachers were asked to check daily the instructional methodologies used in their teaching. They
were also asked to circle the check if a methodology employed during the experiment had not previously been used during the school year. At the conclusion of the experiment the checklist log indicated the frequency and variety of instructional methodology utilization. Trends and patterns were apparent for each teacher as well as for each group.

To facilitate an analysis of the checklist log graphs illustrating the teachers' use of instructional methodologies are presented in Figures 3, 4, and 5. Each graph indicates the degree of teacher methodology use for the three teachers in the group at three times during the experiment. The initial reading was taken during the five day period prior to the commencement of the experiment, the second reading during a five day interval at the

![Figure 3](image)

**Group B--Incidence of Educational Methodology Employment vs. Time**

- This upper cluster represents use of instructional devices and teacher methods
- This lower cluster represents use of instructional devices
Figure 4

Group C—Incidence of Educational Methodology Employment vs. Time

This upper cluster represents use of instructional devices and teacher methods

This lower cluster represents use of instructional devices

Figure 5

Group D—Incidence of Educational Methodology Employment vs. Time

This upper cluster represents use of instructional devices and teacher methods

This lower cluster represents use of instructional devices

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
midway point in the experiment, and the final reading during the
last week of the experiment. These three time periods are desig­
nated t₁, t₂ and t₃ respectively, and are plotted along the abscissas
of the graphs. Numerical scores representing a quantitative measure
of the use of instructional methodologies are plotted along the
ordinates of the graphs. These numerical scores were obtained by
summing the following point designations for methodology employment:

1. Two points were awarded for the first time a specific
methodology was employed during the five day period.

2. One point was awarded for each succeeding time that the
same methodology was employed during the five day period.

3. A five point bonus was awarded for use during the five day
period of a methodology not previously employed by the teacher
during the current school year.

This point system was used so that variety as well as mere quantity
of methodology use was rewarded.

Each graph has two clusters of three lines each. The lower
cluster represents the use of instructional devices during the
experiment while the upper level cluster represents the total of
both instructional devices and teaching methods (instructional
methodologies). Indicated in parentheses behind the upper level
lines is the number of methodologies employed by that teacher during
the experiment for the first time of the current school year.

Care should be taken in the interpretation of the graphs since
they represent readings at only three one-week intervals of the
experiment. This care should especially be exercised with respect to
the initial time interval since the checks here were recorded from
recollection. Nonetheless, the graphs are believed to be indicative
of the relative degree of methodology employment during the experi-
ment.

The graph of group B, the log group, would seem to suggest that
these teachers increased their total methodology use only slightly
during the experiment. The use of instructional devices, however,
was increased considerably. A look at the actual checklist logs
reveals generally a more varied and less routine use of the various
educational methodologies as the experiment progresses. The three
teachers in this group used during the experiment only four method-
ologies not employed previously during the school year.

The graph of group C, the planning group, indicates a decided
increase in both total methodology employment and the use of in-
structional devices. A more scattered rather than linear pattern
of the checklist log further reveals a more varied use of instruc-
tional methodologies. The three teachers in this group used during
the experiment a total of twenty-four methodologies not previously
employed during the school year.

Finally, the graph of group D, the combination group, also
shows a marked increase in total methodology use. Though erratic
in pattern the use of instructional devices also increased slightly.
The three teachers in the combination group employed during the
experiment a total of twenty-one methodologies not previously used
during the school year.
CHAPTER V
SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

This final chapter of the report contains a review of the problem and procedures, a summarization of the major findings, and a presentation of the conclusions. Also included are recommendations for possible use of the data as well as recommendations for further research.

Review of the Problem

The primary purpose of this study was to determine if teachers could modify student perceptions of their teaching performances by employing multiple educational methodologies in their teaching. A secondary objective was to ascertain the relative effectiveness of four experimental treatments in helping a teacher to modify student perceptions of his teaching performance. Five groups of teachers, employing prescribed treatments, were formed in an attempt to gather data pertinent to the objectives of the investigation. Teachers in experimental group A, referred to as the feedback group, received written feedback from students regarding student perceptions of their teaching performances. The treatment for teachers in group B, the log group, was to maintain a checklist log of the various instructional methodologies they employed in their teaching. Teachers in group C, known as the planning group, met weekly with a consultant to make plans for the employment of multiple educational methodologies.
in their teaching. Participants in this group also maintained the checklist log. A combination of the treatments used in groups A, B, and C was utilized by teachers in group D. In addition a fifth group of teachers, group E, was used as a control group.

The criterion measure for obtaining the data was an adapted form of Bryan's Student Opinion Questionnaire (Appendix A).

In order to carry out the objectives of this study four major questions were investigated:

1. Did each of the groups show a significant difference in overall student perception of teacher performance from pre- to post-test measure?

2. Was there a significant difference between groups on the overall posttest measure?

3. Did each of the groups show a significant difference in student perception of the teachers' "variety in teaching" (item nine, SOQ) from pre- to posttest measure?

4. Was there a significant difference between groups for "variety in teaching" (item nine, SOQ) on the posttest measure?

Review of the Procedures Used in the Study

The design of this experiment followed the sequence of pretest--treatment--posttest. As such the experiment paralleled the classical design of research. An adapted form of the SOQ was administered as a pretest and also as a posttest following various experimental treatments during an eight-week interval. The statistical model used to test for significant differences from pre- to posttest measure was
the t-test. To test for significant differences among groups a one-way analysis of variance statistical model was selected as most appropriate.

This field experiment was implemented in three similar contiguous school districts in southeastern Michigan. Fifty-four teachers in the three districts who were currently in their first, second, or third year of teaching and were presently assigned to a fourth, fifth, or sixth-grade self-contained classroom were asked to participate in the initial phase of the study. Following the administration of the adapted SQ pretest fifteen teachers, capable of showing improvement with respect to student perceptions of their teaching performances, were randomly assigned to the five treatment groups. These fifteen teachers, three each in the five groups, then employed the prescribed treatments previously described. At the conclusion of the eight-week treatment period the adapted SQ was administered as a posttest. The pre- and posttest data provided by the 370 students of the fifteen experimental phase teachers were keypunched into data processing cards and a computer program was selected that would yield the necessary statistical information in accord with the design of the study. These data made possible analyses related to four major questions and thirty component questions.

Summary of the Findings

The results of the thirty specific questions studied in this experiment were presented and analyzed in Chapter IV. The findings
related to the four major questions of this investigation are summarized as follows:

1. The results of the investigation indicated that none of the five groups recorded a significant change in overall student perception of teacher performance from pre- to posttest measure.

2. With respect to the relative effectiveness of the four experimental groups in modifying overall student perception of teacher performance the study revealed no significant differences.

3. The investigation revealed several significant changes in students' perceptions of their teachers' "variety in teaching" (item nine, SOQ). The log group, the planning group, and the combination group all showed significant differences at the .001 level from pre- to posttest measure. Neither the feedback group nor the control group showed significant improvements regarding students' perceptions of their teachers' "variety in teaching."

4. Analysis of the data designed to reveal differences in the relative effectiveness of the four experimental groups in modifying students' perceptions of their teachers' "variety in teaching" indicates that differences did exist at the .01 level of significance. The log, planning, and combination groups were all significantly more effective in modifying students' perceptions regarding the teachers' "variety in teaching" than either the feedback or control group. There was no significant difference between the control group and the feedback group. Additionally, no significant differences were found between the log, planning, and combination groups, all of which were successful (.001 level of significance), in effecting a
change from pre- to posttest measure.

5. The data gathered from teacher participants indicate that value was seen in the use of feedback, maintenance of the log, and weekly planning sessions with a person skilled in the use of a variety of instructional methodologies. Feedback was cited as being valuable with respect to an analysis of teacher performance. The checklist log was seen as being beneficial for detecting routine use of instructional methodologies while at the same time challenging the teacher to employ a variety of methods in his teaching. Finally, the planning sessions were most beneficial in assisting teachers to try new and varied methodologies in lesson presentations.

6. An analysis of the checklist log indicates that all groups utilizing the log increased their variety in teaching. All teachers in these groups (B, C, and D) also indicated they had employed more educational methodologies during than prior to the experiment. Teachers planning with a consultant (groups C and D), however, utilized much more frequently instructional methodologies not previously used during the school year.

7. The one-hundred percent participation response of the fifty-four teachers contacted would seem to indicate that upper elementary teachers are eager to participate in studies related to the improvement of the teaching-learning process.

Conclusions

To the extent that the techniques employed may be valid, the following conclusions seem justified:
1. None of the experimental groups was significantly effective in modifying students' overall perceptions of teachers' performances from pre- to posttest measure.

2. No group was significantly more effective than any other group in modifying students' overall perceptions of teachers' performances from pre- to posttest measure.

3. The feedback group treatment was ineffective in helping a teacher to modify student perceptions of his "variety in teaching."

4. The log group, the planning group, and the combination group were all significantly effective in helping a teacher to modify student perceptions of his "variety in teaching."

5. Of the three groups (log, planning, and combination) successful in helping a teacher to modify student perceptions of his "variety in teaching" none was significantly more effective than the other two groups.

6. Teachers in the experimental groups viewed favorably the feedback information, the maintenance of the log, and the weekly planning sessions.

7. The log, planning, and combination groups all showed a trend toward more methodology usage as the experiment progressed.

8. The groups receiving consultative assistance in methodology use (the planning and combination groups) employed far more methodologies not previously used during the school year than the log group.

9. Students are capable of discerning change in a teacher's efforts for improvement and this student perception can be elicited
on a form such as the adapted Student Opinion Questionnaire.

10. Upper elementary teachers were willing and eager participants in the experiment.

In summary the experimental treatments employed seemed to be ineffective in significantly modifying students' overall perceptions of their teachers' performances. Three of the experimental groups (log, planning, and combination), however, appeared to be significantly effective in modifying students' perceptions of their teachers' "variety in teaching."

Recommendations

This section of the report will focus on the proposition of recommendations in two areas: The first will be to suggest possible use of the data by those concerned with the pedagogical dimension of teaching. A second area will be to present implications of this study for further research.

A primary finding of this experiment was that, under the conditions of this experiment, the employment of multiple educational methodologies seemed ineffective in modifying student perception of overall teaching performance. Consequently, personnel in teacher preparation institutions, principals, and curriculum supervisors, should be reluctant to view the employment of multiple educational methodologies as a panacea for curing other classroom ills or improving overall teacher performance as perceived by students.

A second important finding of this study was that three treatment groups were capable of effecting a significant improvement
in student perception of teachers' "variety in teaching." Of the three successful group treatments the log maintenance is obviously the least costly in time and personnel and, therefore, perhaps the most easily implemented. Results of the study would support the contention that by having current teaching staffs maintain the log students' perceptions of a teacher's "variety in teaching" could be expected to improve. The planning group treatment, combining planning with a consultant and log maintenance, also appears to have a decided positive effect on students' perceptions of a teacher's "variety in teaching." Though more costly in time and personnel this treatment has the added feature of being effective in getting teachers to try new methodologies in their teaching. The combination group, utilizing feedback, log maintenance, and planning should also be considered a very definite and effective in-service possibility. No special benefits were found to accrue to the combination group as opposed to the planning group. It is likely, however, that an analysis of the feedback information could provide the readiness and motivation needed for in-service training aimed at the improvement of teacher performances. It is recommended that instruments such as those used in this study could be utilized for the purposes stated above.

Personnel responsible for both pre-service and in-service education might find more viable the procedure of placing a concerted effort on a given area of teacher performance with the goal in mind of improving that one facet, rather than taking a more global approach toward improvement. Improvement in other areas most likely would
require additional concerted efforts again aimed specifically at improving that one facet of the total performance.

Additionally, it would appear that if teachers are to attempt the employment of new methodologies the suggestions and reinforcement of consultative type personnel would be beneficial.

With regard to the implications of this study for further research it is apparent that studies designed to go beyond the scope of the present investigation would be advantageous in the quest for improving teacher performance.

It would seem to be of value to repeat the basic thrust of this experiment on a larger scale. Previously mentioned as a limitation of the study was the small number of experimental group teachers. A team of researchers might find it feasible to increase this number and make broader generalizations than were possible in this experiment. A further recommendation would be to partition teachers on the quality of their initial student-perceived performances. Such a study might also be conducted over a longer period of time than that allotted for this experiment. Finally, if a future study were carried out on a larger scale and on a longitudinal basis it might be possible to determine the net improvement of a teacher apart from the lasting initial perceptions of students. For example, if the image is "stubbornly stable," teachers desirous of improvement first find it necessary to erase initial perceptions of a given group of students. If, however, on a larger scale, teachers were rated midway through a given year, participated in a concerted effort to change a
given facet of their performance, and then were rated again by another class midway through a second year, it would not be necessary to erase the opinions of the first group of students in computing the net improvement of a teacher's performance.

Although the present study showed that with a concerted effort teachers were able to improve student perceptions of their "variety in teaching," it also revealed that a manipulation of this one item did not improve student perceptions of the teachers' overall performances. Two questions might merit additional study: Is it possible to improve on other commonly accepted criteria of teacher effectiveness with a concerted effort and does a manipulation of one such criterion hold a key for improving student perceptions of the teachers' overall performances?

One obvious concern generated by this experiment is the inconclusiveness of the benefit of feedback information. Many studies previously cited in the related literature section of this report have espoused its effectiveness. Bryan\(^1\) tempered his findings somewhat by referring to the teacher's image as "stubbornly stable." Lauroesch\(^2\) found, as this study did, that teachers receiving feedback over a short period of time actually decreased in student-perceived performance.

---

\(^1\) Bryan, Roy C., "The Teacher's Image is Stubbornly Stable." The Clearing House, XL (April 1966), 459.

\(^2\) Lauroesch, op. cit., p. 2.
effectiveness. Is it possible that teacher quality is a factor in improving teacher performance? In short, is there a difference in the relative effectiveness of feedback in modifying the performance of above average, average, or below average teachers as perceived by students?

Another consideration regarding feedback deals with the time element. Is eight weeks following the reception of feedback information too short a time to register significant changes in student perception?

The enthusiasm displayed by both teachers and administrators for this study at the upper elementary level hopefully is indicative of possible receptivity for future studies of this kind. A meaningful comparison could be gained by a replication of this experiment at the senior high level.
BIBLIOGRAPHY


Brady, Elizabeth H. and Richardson, Sybil K., "How Do Children View the Teacher?" Childhood Education, XXXVI (November 1959), 111-2.


Callahan, Sterling G., "Is Teacher Rating by Students a Sound Practice"? *School and Society*, LXIX (February 1949), 98-100.


Cosgrove, Don J., "Diagnostic Rating of Teacher Performance." *Journal of Educational Psychology*, L (October 1959), 200-4.


Mason, Barbara T., "Supervisor or Curriculum Specialist." Educational Leadership, XXVII (January 1970), 401-3.


Raths, Louis, "What is a Good Teacher"? Childhood Education, XL (May 1964), 451-6.


Tedesco, Phyliss Reynolds, "An 'A' for Teacher." The Instructor, LXIX (October 1959), 89.


Tuckman, B. W. and Oliver, W. F., "Effectiveness of Feedback to Teachers as a Function of Source." Journal of Educational Psychology, LIX (August 1968), 297-301.


APPENDIXES
APPENDIX A

Student Opinion Questionnaire

Follow along with me as I read the instructions aloud. There should be no talking. This is not a test. There are no right or wrong answers. Please answer the questions honestly. Do not give your name. Neither your teacher nor anyone else at your school will ever see your answers.

All your answers will be sealed in this envelope and taken for scoring. Your teacher will receive a summary of the answers by the students in your class. Your teacher wants you to answer each question honestly. The summary will help her to know what the class likes and dislikes about her teaching. This could help her become a better teacher.

CIRCLE THE ANSWER WHICH BEST GRADES YOUR TEACHER’S:

Example A. APPEARANCE: (Does your teacher wear nice clothes?)

<table>
<thead>
<tr>
<th>POOR</th>
<th>FAIR</th>
<th>AVERAGE</th>
<th>GOOD</th>
<th>EXCELLENT</th>
</tr>
</thead>
</table>

Example B. PENMANSHIP: (How well does your teacher write on the board?)

<table>
<thead>
<tr>
<th>POOR</th>
<th>FAIR</th>
<th>AVERAGE</th>
<th>GOOD</th>
<th>EXCELLENT</th>
</tr>
</thead>
</table>

1. KNOWLEDGE: (How well does your teacher understand the lessons she teaches to you?)

<table>
<thead>
<tr>
<th>POOR</th>
<th>FAIR</th>
<th>AVERAGE</th>
<th>GOOD</th>
<th>EXCELLENT</th>
</tr>
</thead>
</table>

2. EXPLANATIONS: (How well does your teacher explain the lessons to the class?)

<table>
<thead>
<tr>
<th>POOR</th>
<th>FAIR</th>
<th>AVERAGE</th>
<th>GOOD</th>
<th>EXCELLENT</th>
</tr>
</thead>
</table>

3. FAIRNESS: (Is your teacher fair with all pupils?)

<table>
<thead>
<tr>
<th>POOR</th>
<th>FAIR</th>
<th>AVERAGE</th>
<th>GOOD</th>
<th>EXCELLENT</th>
</tr>
</thead>
</table>

4. CONTROL: (How well do pupils behave for your teacher?)

<table>
<thead>
<tr>
<th>POOR</th>
<th>FAIR</th>
<th>AVERAGE</th>
<th>GOOD</th>
<th>EXCELLENT</th>
</tr>
</thead>
</table>

5. SUCCESS IN KEEPING YOUR ATTENTION: (Does your teacher make the lessons interesting?)

<table>
<thead>
<tr>
<th>POOR</th>
<th>FAIR</th>
<th>AVERAGE</th>
<th>GOOD</th>
<th>EXCELLENT</th>
</tr>
</thead>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
6. **ENTHUSIASM:** (Does your teacher like to teach?)

<table>
<thead>
<tr>
<th></th>
<th>POOR</th>
<th>FAIR</th>
<th>AVERAGE</th>
<th>GOOD</th>
<th>EXCELLENT</th>
</tr>
</thead>
</table>

7. **CONSIDERATION:** (Is your teacher kind toward all pupils in your class?)

<table>
<thead>
<tr>
<th></th>
<th>POOR</th>
<th>FAIR</th>
<th>AVERAGE</th>
<th>GOOD</th>
<th>EXCELLENT</th>
</tr>
</thead>
</table>

8. **ATTITUDE TOWARD STUDENTS:** (Does your teacher like you?)

<table>
<thead>
<tr>
<th></th>
<th>POOR</th>
<th>FAIR</th>
<th>AVERAGE</th>
<th>GOOD</th>
<th>EXCELLENT</th>
</tr>
</thead>
</table>

9. **VARIETY IN TEACHING:** (Does your teacher usually teach the same way day after day or does she often try different ways to present lessons?) (motion pictures, tape recorders, record players, charts, overhead projectors, student reports, small group projects, guest speakers, demonstrations, field trips, etc.)

<table>
<thead>
<tr>
<th></th>
<th>POOR</th>
<th>FAIR</th>
<th>AVERAGE</th>
<th>GOOD</th>
<th>EXCELLENT</th>
</tr>
</thead>
</table>

10. **ENCOURAGEMENT OF STUDENT PARTICIPATION:** (Does your teacher like students to ask questions?)

<table>
<thead>
<tr>
<th></th>
<th>POOR</th>
<th>FAIR</th>
<th>AVERAGE</th>
<th>GOOD</th>
<th>EXCELLENT</th>
</tr>
</thead>
</table>

11. **SENSE OF HUMOR:** (Does your teacher laugh with the class?)

<table>
<thead>
<tr>
<th></th>
<th>POOR</th>
<th>FAIR</th>
<th>AVERAGE</th>
<th>GOOD</th>
<th>EXCELLENT</th>
</tr>
</thead>
</table>

12. **PLANNING:** (Does your teacher make good plans for the use of class time so that little time is wasted?)

<table>
<thead>
<tr>
<th></th>
<th>POOR</th>
<th>FAIR</th>
<th>AVERAGE</th>
<th>GOOD</th>
<th>EXCELLENT</th>
</tr>
</thead>
</table>

13. **ASSIGNMENTS:** (Does your teacher make your homework enjoyable?)

<table>
<thead>
<tr>
<th></th>
<th>POOR</th>
<th>FAIR</th>
<th>AVERAGE</th>
<th>GOOD</th>
<th>EXCELLENT</th>
</tr>
</thead>
</table>

A. **What is the ONE thing that you like MOST about your teacher?**

B. **What is the ONE thing that you like LEAST about your teacher?**
## APPENDIX B

### Teacher Profile

**Teacher** ____________  **Grade** ____________  **No. of Students** ____________  **Date** ____________

<table>
<thead>
<tr>
<th>Scale Steps</th>
<th>Scale Average</th>
<th>ITEMS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EXCELLENT</strong></td>
<td>5.0</td>
<td>1 2 3 4 5 6 7 8 9 10 11 12 13 14</td>
</tr>
<tr>
<td></td>
<td>4.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.1</td>
<td></td>
</tr>
<tr>
<td><strong>GOOD</strong></td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.1</td>
<td></td>
</tr>
<tr>
<td><strong>AVERAGE</strong></td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td><strong>FAIR</strong></td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.3</td>
<td></td>
</tr>
</tbody>
</table>
## Appendix B (continued)

<table>
<thead>
<tr>
<th>scale steps</th>
<th>scale average</th>
<th>ITEMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2</td>
<td></td>
<td>1 2 3 4 5 6 7 8 9 10 11 12 13 14</td>
</tr>
<tr>
<td>1.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>POOR</td>
<td>1.0</td>
<td></td>
</tr>
</tbody>
</table>

**Key to Items**

1. Knowledge  
2. Explanations  
3. Fairness  
4. Control  
5. Interest  
6. Enthusiasm  
7. Consideration  
8. Attitude Toward Students  
9. Variety  
10. Student Participation  
11. Sense of Humor  
12. Planning  
13. Assignments  
14. Mean of Averages 1-13

A. Strengths listed by a significant number of students:

B. Weaknesses listed by a significant number of students:
APPENDIX C

Tabular Summation

<table>
<thead>
<tr>
<th>SCALE STEPS</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTALS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX D

Feedback

Enclosed is a compilation of student responses to the questionnaire given recently to your class. At the top of the page you will notice a graph. This graph is an average of student responses to questions 1-13 of the Student Opinion Questionnaire. A key to the items in the questionnaire and an average for each appears at the bottom of the page. Note what your class perceives to be your strengths and weaknesses, paying particular attention to those items which appear too low. Hopefully this information will suggest areas in which you may wish to make a special effort to change your students' perceptions.
APPENDIX E

Checklist Log

Instructions

You have been selected to participate in an additional phase of the total field experiment previously described. Your cooperation is deeply appreciated and an essential and significant part of the experimental study.

During the eight weeks of the experiment you are asked to maintain a checklist or "log" of the various instructional methods you employ. The methods and devices listed on the attached sheet are commonly accepted in the literature and in practice as appropriate instructional techniques.

Your specific instructions are:

1) At the end of each day check the methods and or devices used during that day. (Only one check per item per day, even though used more than once)

2) It is possible that an instructional device (Group I) would be used in conjunction with a teaching method (Group II). For example, the overhead projector might be used in a lecture type presentation. In instances of this type a check should be made for the overhead projector and for the lecture.

3) The possibility also exists that more than one teaching method (Group II) might be employed during a given learning experience. In this case check only the method which is predominant and do not check the method which is employed incidentally. For example, a teacher conducts a discussion type presentation. During the course of the discussion a few questions are asked. The discussion is the predominant method employed and this category should be checked. The questions are incidental to the discussion and the question-answer category should not be checked.

To aid you in making consistent and mutually exclusive checks the following operational descriptions are offered for items which may not be self-explanatory:
DISCUSSION: Group interaction characterized by a free exchange of views and ideas by participants. Oral two-way communication during which learners examine, consider, and compare facts and reasoning. Sometimes called group thinking aloud.

DRILL: Commonly referred to as recitation or practice usually for the purpose of automatic recall. Examples would be spelling, routine math problems, one word question-answer, workbooks, etc. Oftentimes follows the pattern of assignment - study - report.

INDEPENDENT STUDY: A major planned learning experience in which a student can proceed in accord with his own initiative and ability using the teacher as a resource. Not the daily, routine, unplanned type of study or reading a student often pursues while waiting for classmates to finish assigned work.

LABORATORY: Direct experience with materials pertinent to the area of study. Usually associated with science lessons, but not necessarily so.

LECTURE: Oral one-way communication, primarily teacher talk. An exposition of facts or knowledge by a teacher.

PROBLEM SOLVING: Closely allied with learning by discovery and the scientific method. It involves the thought process of logical and critical thinking that results from a doubt, a perplexity, a problem or an observation. The recognition of a problem is followed sequentially by an inspection of possible solutions, the employment of a given solution, and the evaluation of its success in solving the problem.

QUESTION-ANSWER: Open-ended critical appraisal type of questions. Frequently used in a well-planned sequence to develop major concepts. Not one word answer type of questions.
## MULTIPLE INSTRUCTIONAL METHODOLOGIES

### I. INSTRUCTIONAL DEVICES

<table>
<thead>
<tr>
<th>Method Description</th>
<th>Date Employed</th>
</tr>
</thead>
<tbody>
<tr>
<td>A) Educational games</td>
<td></td>
</tr>
<tr>
<td>B) 8mm cartridge projector</td>
<td></td>
</tr>
<tr>
<td>C) Filmstrip projector</td>
<td></td>
</tr>
<tr>
<td>D) Motion picture projector</td>
<td></td>
</tr>
<tr>
<td>E) Opaque projector</td>
<td></td>
</tr>
<tr>
<td>F) Overhead projector</td>
<td></td>
</tr>
<tr>
<td>G) Programmed instruction</td>
<td></td>
</tr>
<tr>
<td>H) Radio</td>
<td></td>
</tr>
<tr>
<td>I) Record player</td>
<td></td>
</tr>
<tr>
<td>J) Slide projector</td>
<td></td>
</tr>
<tr>
<td>K) Tape recorder</td>
<td></td>
</tr>
<tr>
<td>L) Television</td>
<td></td>
</tr>
<tr>
<td>M) Chart, globe, model, etc.</td>
<td></td>
</tr>
<tr>
<td>N) Other (specify)</td>
<td></td>
</tr>
</tbody>
</table>

### II. TEACHING METHODS

<table>
<thead>
<tr>
<th>Method Description</th>
<th>Date Employed</th>
</tr>
</thead>
<tbody>
<tr>
<td>A) Demonstration</td>
<td></td>
</tr>
<tr>
<td>B) Discussion</td>
<td></td>
</tr>
<tr>
<td>C) Drill (recitation)</td>
<td></td>
</tr>
<tr>
<td>D) Evaluation -- Review</td>
<td></td>
</tr>
</tbody>
</table>
|\[ Test \]
|\[ Quiz \]
| E) Field Trip |             |
| F) Group methods -- Debate |    |
|\[ Panel \]
| \[ Project \]
| G) Independent study |         |
| H) Individual report |            |
| I) Laboratory |             |
| J) Lecture |                |
| K) Guest speaker |             |
| L) Problem solving (discovery) |   |
| M) Question-Answer |           |
| N) Role playing |              |
| O) Teacher-pupil planning |         |
| P) Team teaching (ad hoc) |        |
| Q) Other (specify) |           |

---

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Fellow Educator:

This letter is written to introduce to you Mr. James Bultman. Jim is currently a Mott Doctoral Fellow working thru Western Michigan University. Previously he has been a teacher, principal and college professor.

We have reviewed the research study proposed by Mr. Bultman and believe it to be significant and noteworthy for us as educators. Jim assures us that the anonymity of participants and their individual results will be honored.

Because of the experimental nature of the study you should understand that it is not possible to reveal certain parts of the study at the outset. It is not anticipated that participation in the study will in any way be an imposition on your time.

In that education is in need of enlightening research studies and because the proposed research is legitimate and worthy of our participation you are urged to give it your careful consideration.

Sincerely,

Superintendent of Schools
APPENDIX G

Computer Program

Hope College Statistical Programs Manual

Variance Analysis

MNSTR: One-Way Analysis of Variance with Within-Group Correlations and Between-Group t-Tests

1. GENERAL DESCRIPTION

a. This program simultaneously computes an F-statistic for a One-Way Analysis of Variance between groups for each variable. Also included in the output are distribution statistics, within-group correlations, and between-group t-values.

b. The output consists of (for each variable):

For each group and total group
1) Mean
2) Standard deviation
3) Estimate of population standard deviation
4) Estimate of population variance
5) Standard Error of the Mean
6) Correlation Matrix
7) t-test of difference from zero for the Mean

For total group only
8) Pooled Within-Group Correlation Matrix

Analyses between groups
9) Analysis of Variance table for each variable consisting of source, sum of squares, degrees of freedom, mean squares, and F-ratio.
10) t-test between each group for each variable using the common mean square as the error term.
11) t-test between each group for each variable using separate error terms; the error term being computed by using only the two groups in question.

c. Limitations:

1) The number of variables must not exceed 25.
2) The number of groups must not exceed 32,767.
3) The number of subjects per group must not exceed 32,767.
4) If the number of groups exceeds 25 the between group t-tests will not be printed.

2. Authors: Dee Norton, Ph.D.
   Department of Psychology
   State University of Iowa
   and
   Bill Snider, Ph.D.
   University Computer Center
   State University of Iowa
   (Modified for use with the IBM 1130 by George Bishop, Hope College)
## APPENDIX H

### Student Perception Averages for All Questionnaire Items, Pre and Post

<table>
<thead>
<tr>
<th>Question</th>
<th>Teacher</th>
<th>Group A</th>
<th>Group B</th>
<th>Group C</th>
<th>Group D</th>
<th>Group E</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre</td>
<td>1</td>
<td>4.3</td>
<td>4.0</td>
<td>3.7</td>
<td>3.9</td>
<td>3.9</td>
</tr>
<tr>
<td>post</td>
<td></td>
<td>4.3</td>
<td>4.5</td>
<td>3.6</td>
<td>3.9</td>
<td>3.8</td>
</tr>
<tr>
<td>post</td>
<td>2</td>
<td>4.0</td>
<td>3.7</td>
<td>4.1</td>
<td>3.8</td>
<td>3.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.8</td>
<td>4.5</td>
<td>4.0</td>
<td>3.6</td>
<td>3.3</td>
</tr>
<tr>
<td>post</td>
<td>3</td>
<td>4.0</td>
<td>4.3</td>
<td>3.6</td>
<td>3.7</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.5</td>
<td>3.8</td>
<td>4.0</td>
<td>3.4</td>
<td>4.0</td>
</tr>
<tr>
<td>post</td>
<td>1</td>
<td>4.0</td>
<td>3.8</td>
<td>2.9</td>
<td>3.8</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.0</td>
<td>3.7</td>
<td>2.6</td>
<td>4.0</td>
<td>3.8</td>
</tr>
<tr>
<td>post</td>
<td>2</td>
<td>3.8</td>
<td>4.0</td>
<td>3.8</td>
<td>3.8</td>
<td>3.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.6</td>
<td>4.0</td>
<td>3.7</td>
<td>3.7</td>
<td>3.3</td>
</tr>
<tr>
<td>post</td>
<td>3</td>
<td>3.9</td>
<td>4.0</td>
<td>3.5</td>
<td>3.9</td>
<td>3.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.8</td>
<td>3.5</td>
<td>3.6</td>
<td>3.1</td>
<td>3.7</td>
</tr>
<tr>
<td>post</td>
<td>1</td>
<td>3.3</td>
<td>3.3</td>
<td>3.5</td>
<td>3.3</td>
<td>3.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.9</td>
<td>3.6</td>
<td>3.1</td>
<td>3.3</td>
<td>3.4</td>
</tr>
<tr>
<td>post</td>
<td>2</td>
<td>3.4</td>
<td>3.8</td>
<td>3.5</td>
<td>3.7</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.9</td>
<td>4.1</td>
<td>3.2</td>
<td>2.8</td>
<td>2.7</td>
</tr>
<tr>
<td>post</td>
<td>3</td>
<td>3.7</td>
<td>3.7</td>
<td>3.1</td>
<td>2.1</td>
<td>3.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.5</td>
<td>2.3</td>
<td>3.0</td>
<td>1.8</td>
<td>3.6</td>
</tr>
<tr>
<td>post</td>
<td>1</td>
<td>3.5</td>
<td>2.9</td>
<td>2.3</td>
<td>3.3</td>
<td>3.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.0</td>
<td>3.2</td>
<td>2.3</td>
<td>3.2</td>
<td>3.6</td>
</tr>
<tr>
<td>post</td>
<td>2</td>
<td>3.0</td>
<td>3.2</td>
<td>2.9</td>
<td>1.9</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.2</td>
<td>3.0</td>
<td>3.2</td>
<td>2.4</td>
<td>2.7</td>
</tr>
<tr>
<td>post</td>
<td>3</td>
<td>3.0</td>
<td>3.0</td>
<td>2.6</td>
<td>2.9</td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.5</td>
<td>2.6</td>
<td>3.6</td>
<td>2.5</td>
<td>2.4</td>
</tr>
<tr>
<td>post</td>
<td>1</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td>3.5</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.8</td>
<td>4.1</td>
<td>2.7</td>
<td>3.8</td>
<td>3.3</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
<table>
<thead>
<tr>
<th>Question</th>
<th>Teacher</th>
<th>Group A</th>
<th>Group B</th>
<th>Group C</th>
<th>Group D</th>
<th>Group E</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>pre</td>
<td>3.3</td>
<td>3.7</td>
<td>3.5</td>
<td>3.0</td>
<td>3.7</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td>3.2</td>
<td>4.0</td>
<td>3.8</td>
<td>2.9</td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td>pre</td>
<td>3.5</td>
<td>2.9</td>
<td>3.8</td>
<td>3.2</td>
<td>3.7</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td>2.5</td>
<td>2.6</td>
<td>3.6</td>
<td>2.6</td>
<td>3.7</td>
</tr>
<tr>
<td>6</td>
<td>pre</td>
<td>4.4</td>
<td>4.0</td>
<td>4.0</td>
<td>4.6</td>
<td>3.8</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td>4.1</td>
<td>4.0</td>
<td>3.8</td>
<td>3.8</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>pre</td>
<td>3.7</td>
<td>4.3</td>
<td>4.0</td>
<td>3.7</td>
<td>3.8</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td>3.9</td>
<td>4.3</td>
<td>3.7</td>
<td>3.5</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>pre</td>
<td>3.8</td>
<td>4.2</td>
<td>4.0</td>
<td>3.9</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td>3.0</td>
<td>3.8</td>
<td>4.2</td>
<td>3.6</td>
<td>4.2</td>
</tr>
<tr>
<td>7</td>
<td>pre</td>
<td>3.2</td>
<td>3.3</td>
<td>3.8</td>
<td>3.8</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td>3.0</td>
<td>3.9</td>
<td>3.4</td>
<td>3.6</td>
<td>3.4</td>
</tr>
<tr>
<td>8</td>
<td>pre</td>
<td>3.4</td>
<td>4.0</td>
<td>3.7</td>
<td>3.3</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td>3.2</td>
<td>3.8</td>
<td>3.5</td>
<td>2.8</td>
<td>2.9</td>
</tr>
<tr>
<td></td>
<td>pre</td>
<td>3.5</td>
<td>4.1</td>
<td>3.0</td>
<td>3.2</td>
<td>3.7</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td>3.0</td>
<td>3.4</td>
<td>3.2</td>
<td>2.6</td>
<td>3.8</td>
</tr>
<tr>
<td>9</td>
<td>pre</td>
<td>3.6</td>
<td>3.7</td>
<td>3.3</td>
<td>3.1</td>
<td>3.8</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td>3.2</td>
<td>3.2</td>
<td>3.4</td>
<td>3.3</td>
<td>4.3</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
<table>
<thead>
<tr>
<th>Question</th>
<th>Teacher</th>
<th>Group A</th>
<th>Group B</th>
<th>Group C</th>
<th>Group D</th>
<th>Group E</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>pre</td>
<td>4.2</td>
<td>3.0</td>
<td>3.8</td>
<td>3.4</td>
<td>3.8</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td>4.3</td>
<td>2.9</td>
<td>3.9</td>
<td>3.8</td>
<td>3.5</td>
</tr>
<tr>
<td>11</td>
<td>pre</td>
<td>3.5</td>
<td>4.3</td>
<td>4.1</td>
<td>3.2</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td>3.2</td>
<td>4.1</td>
<td>4.7</td>
<td>3.1</td>
<td>4.2</td>
</tr>
<tr>
<td>12</td>
<td>pre</td>
<td>4.1</td>
<td>3.7</td>
<td>3.5</td>
<td>4.5</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td>4.3</td>
<td>4.1</td>
<td>3.5</td>
<td>4.0</td>
<td>3.5</td>
</tr>
<tr>
<td>13</td>
<td>pre</td>
<td>2.4</td>
<td>3.8</td>
<td>4.2</td>
<td>3.1</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td>2.0</td>
<td>3.0</td>
<td>2.6</td>
<td>1.7</td>
<td>3.6</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
### APPENDIX H (continued)

<table>
<thead>
<tr>
<th>Question</th>
<th>Teacher</th>
<th>Group A</th>
<th>Group B</th>
<th>Group C</th>
<th>Group D</th>
<th>Group E</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pre</td>
<td>3.5</td>
<td>3.7</td>
<td>3.7</td>
<td>3.3</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td>3.5</td>
<td>3.9</td>
<td>3.6</td>
<td>3.2</td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td>pre</td>
<td>3.4</td>
<td>3.6</td>
<td>3.3</td>
<td>3.0</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td>3.0</td>
<td>3.4</td>
<td>3.7</td>
<td>3.0</td>
<td>3.6</td>
</tr>
</tbody>
</table>

APPENDIX I

Feedback Group Open-Ended Response

Please comment about the effect (if any) of the experimental "treatment" on your teaching during the course of the experiment.

Feedback:

General comments you wish to make about the field experiment:
APPENDIX J

Log Group Open-Ended Response

Please comment about the effect (if any) of the experimental "treatment" on your teaching during the course of the experiment.

Checklist log:

General comments you wish to make about the field experiment:

Did you, during the course of the experiment, employ a greater variety of instructional methodologies than prior to the experiment? (YES or NO)
Planning Group Open-Ended Response

Please comment about the effect (if any) of the experimental "treatment" on your teaching during the course of the experiment.

Checklist log:

Weekly planning sessions:

General comments you wish to make about the field experiment:

Did you, during the course of the experiment, employ a greater variety of instructional methodologies than prior to the experiment? (YES or NO)
APPENDIX L

Combination Group Open-Ended Response

Please comment about the effect (if any) of the experimental "treatment" on your teaching during the course of the field experiment.

Feedback:

Checklist log:

Weekly planning sessions:

General comments you wish to make about the field experiment:

Did you, during the course of the experiment, employ a greater variety of instructional methodologies than prior to the experiment? (YES or NO)

156
**APPENDIX M**

Teacher Self-perception Averages for All Questionnaire Items, Pre and Post

<table>
<thead>
<tr>
<th>Question</th>
<th>Teacher</th>
<th>Group A</th>
<th>Group B</th>
<th>Group C</th>
<th>Group D</th>
<th>Group E</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>pre</td>
<td>post</td>
<td>pre</td>
<td>post</td>
<td>pre</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4.0</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.0</td>
<td>4.0</td>
<td>5.0</td>
<td>3.0</td>
<td>4.0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>4.0</td>
<td>4.0</td>
<td>2.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3.0</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.0</td>
<td>3.0</td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.0</td>
<td>4.0</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2.0</td>
<td>3.0</td>
<td>5.0</td>
<td>4.0</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.0</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
<td>2.0</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>5.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3.0</td>
<td>5.0</td>
<td>4.0</td>
<td>3.0</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.0</td>
<td>5.0</td>
<td>4.0</td>
<td>3.0</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2.0</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.0</td>
<td>3.0</td>
<td>5.0</td>
<td>4.0</td>
<td>4.0</td>
</tr>
</tbody>
</table>

157

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
<table>
<thead>
<tr>
<th>Question</th>
<th>Teacher</th>
<th>Group A</th>
<th>Group B</th>
<th>Group C</th>
<th>Group D</th>
<th>Group E</th>
</tr>
</thead>
<tbody>
<tr>
<td>pre 1</td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
<td>2.0</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>post</td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>pre 2</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>post</td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>pre 3</td>
<td>3.0</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>post</td>
<td>3.0</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>pre 1</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td>2.0</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>post</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
<td>5.0</td>
<td></td>
</tr>
<tr>
<td>pre 2</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>post</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>pre 3</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>2.0</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>post</td>
<td>4.0</td>
<td>4.0</td>
<td>5.0</td>
<td>4.0</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>pre 1</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>post</td>
<td>4.0</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>pre 2</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
<td>5.0</td>
<td></td>
</tr>
<tr>
<td>post</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>pre 3</td>
<td>3.0</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>post</td>
<td>3.0</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>pre 1</td>
<td>4.0</td>
<td>5.0</td>
<td>4.0</td>
<td>3.0</td>
<td>5.0</td>
<td></td>
</tr>
<tr>
<td>post</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>pre 2</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>post</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>pre 3</td>
<td>3.0</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>post</td>
<td>3.0</td>
<td>3.0</td>
<td>5.0</td>
<td>4.0</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>pre 1</td>
<td>3.0</td>
<td>5.0</td>
<td>2.0</td>
<td>1.0</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>post</td>
<td>2.0</td>
<td>3.0</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
<td></td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
<table>
<thead>
<tr>
<th>Question</th>
<th>Teacher</th>
<th>Group A</th>
<th>Group B</th>
<th>Group C</th>
<th>Group D</th>
<th>Group E</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>pre</td>
<td>2</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td></td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>pre</td>
<td>3</td>
<td>3.0</td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td></td>
<td>4.0</td>
<td>3.0</td>
<td>5.0</td>
<td>5.0</td>
</tr>
<tr>
<td>10</td>
<td>pre</td>
<td>1</td>
<td>4.0</td>
<td>5.0</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td></td>
<td>4.0</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>pre</td>
<td>2</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td></td>
<td>5.0</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>pre</td>
<td>3</td>
<td>4.0</td>
<td>4.0</td>
<td>5.0</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td></td>
<td>3.0</td>
<td>4.0</td>
<td>5.0</td>
<td>5.0</td>
</tr>
<tr>
<td>11</td>
<td>pre</td>
<td>1</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td></td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>pre</td>
<td>2</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td></td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>pre</td>
<td>3</td>
<td>4.0</td>
<td>4.0</td>
<td>5.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td></td>
<td>3.0</td>
<td>4.0</td>
<td>5.0</td>
<td>5.0</td>
</tr>
<tr>
<td>12</td>
<td>pre</td>
<td>1</td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td></td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>pre</td>
<td>2</td>
<td>3.0</td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td></td>
<td>3.0</td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>pre</td>
<td>3</td>
<td>3.0</td>
<td>3.0</td>
<td>4.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td></td>
<td>3.0</td>
<td>3.0</td>
<td>5.0</td>
<td>4.0</td>
</tr>
<tr>
<td>13</td>
<td>pre</td>
<td>1</td>
<td>4.0</td>
<td>2.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td></td>
<td>4.0</td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>pre</td>
<td>2</td>
<td>2.0</td>
<td>4.0</td>
<td>3.0</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td></td>
<td>4.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>pre</td>
<td>3</td>
<td>3.0</td>
<td>3.0</td>
<td>4.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td></td>
<td>2.0</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td>Question</td>
<td>Teacher</td>
<td>Group A</td>
<td>Group B</td>
<td>Group C</td>
<td>Group D</td>
<td>Group E</td>
</tr>
<tr>
<td>----------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>pre</td>
<td>4.0</td>
<td>3.8</td>
<td>3.2</td>
<td>2.9</td>
<td>3.8</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td>3.8</td>
<td>3.4</td>
<td>3.3</td>
<td>3.5</td>
<td>3.7</td>
</tr>
<tr>
<td>T</td>
<td>pre</td>
<td>3.6</td>
<td>4.1</td>
<td>3.4</td>
<td>3.4</td>
<td>3.7</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td>3.8</td>
<td>3.9</td>
<td>3.5</td>
<td>3.6</td>
<td>3.4</td>
</tr>
<tr>
<td></td>
<td>pre</td>
<td>3.2</td>
<td>3.3</td>
<td>4.2</td>
<td>3.7</td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td>post</td>
<td>3.2</td>
<td>3.2</td>
<td>4.5</td>
<td>4.1</td>
<td>3.3</td>
</tr>
</tbody>
</table>