

A PILOT STUDY TESTING
EFFECTS OF FEEDBACK ON TEACHING BEHAVIOR
AS MEASURED BY STUDENT RATINGS AT
NEW RIVER COMMUNITY COLLEGE,
DUBLIN, VIRGINIA

by

Norman H. Scott

Dissertation submitted to the Graduate Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of
DOCTOR OF EDUCATION

in

Vocational and Technical Education

APPROVED:

Dr. Dewey A. Adams, Chairman

Dr. David M. Moore

Dr. Samuel D. Morgan

Dr. N. Alan Sheppard

Dr. Harold W. Stubblefield

April, 1976

Blacksburg, Virginia

ACKNOWLEDGMENTS

The writer wishes to acknowledge assistance of those who have taken an active interest in the development and preparation of this study.

Special appreciation is expressed to Dr. Dewey A. Adams, Chairman of the writer's advisory committee. Appreciation is also expressed to Dr. David M. Moore, Dr. Samuel D. Morgan, Dr. N. Alan Sheppard, and Dr. Harold W. Stubblefield, members of the committee.

Gratitude is expressed for special assistance from Mr. John Endall, Dr. Dennis E. Hinkle, and Dr. Clyde Y. Kramer in preparing the statistical program used in this study and to Sybil O. Sutphen for her assistance in collecting data.

The writer's gratitude is also expressed to his parents, Beth and Norman Scott, for their encouragement and support, and to his wife, Grace, and son, Norman, for their unlimited patience and sacrifice while this study was undertaken.

Special thanks go to Dr. W. Robert Sullins, President, and the faculty of New River Community College, for their cooperation with and assistance to the writer in this study.

7

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS	ii
LIST OF TABLES	vii
Chapter	
I. THE PROBLEM	1
Description of the Problem Situation	1
Statement of the Problem	2
Objectives of the Study	3
Assumptions	3
Limitations of the Study	4
Definition of Terms	5
Significance of the Study	5
Organization of the Study	7
II. REVIEW OF THE LITERATURE	8
Need for Self-Evaluation	8
Summary	14
Use of Feedback	15
Summary	21
Use of Closed Circuit Television	22
Summary	32
A Combination of Closed Circuit Television and Other Feedback Devices	33
Summary	37

Chapter	Page
Student Feedback	38
Summary	42
Summary and Theoretical Framework	43
Summary	43
Theoretical Framework	44
Theoretical Schema	45
Statement of Hypotheses	45
III. RESEARCH METHODOLOGY	47
Introduction	47
Population	47
Research Design	48
Treatments	49
Controls for Threats to Internal Validity .	51
Controls for Threats to External Validity .	52
Sample Selection	53
Instrumentation and Data Collection	55
Instrumentation	55
Purpose of the Instrument	55
Reliability of the Instrument	58
Validity of the Instrument	58
Item Omission	58
Data Collection	59
Statistical Analysis	60
Treatment of Data	60
Computation of Gain Scores	60

Chapter	Page
Test of the Hypotheses	61
Relationship of Demographic Characteristics. .	62
Time Schedule	63
Summary	64
IV. PRESENTATION AND ANALYSES OF DATA	66
Collection of Data	66
Computation of Gain Scores	66
Test of Hypothesis One	67
Test of Hypothesis Two	67
Relationship of Demographic Characteristics to Teaching Behavior	71
Teaching Area	71
Teaching Experience	71
Age	74
Educational Level	74
Occupational Experience Other Than Teaching	74
Summary	74
Test of Hypothesis One	74
Test of Hypothesis Two	74
Relationship of Demographic Characteristics to Teaching Behavior	78
V. SUMMARY, CONCLUSIONS AND RECOMMENDATIONS	79
Summary	79
Statement of the Problem	79
Statement of Hypotheses	80

Chapter	Page
Research Methodology	80
Test of the Hypotheses	81
Relationship of Demographic Characteristics To Teaching Behavior	81
Results	82
Discussion of Results	83
Recommendations	85
REFERENCES	88
APPENDIX	96
VITA	102

LIST OF TABLES

Table		Page
1.	Diagram of Research Design	50
2.	Number of Population Subjects According to Teaching Area Which Fall into Each of the Three Subdivisions of Teaching Experience	56
3.	Number of Sample Subjects For Treatment and Control Groups	57
4.	Multivariate T^2 Test for Hypothesis One	68
5.	Multivariate Analysis of Variance for Hypothesis Two	69
6.	Linear Contrasts by Pairs of Treatments for Hypothesis Two	70
7.	Multivariate Analysis of Variance Illustrating Relationship Between Teaching Area and Instrument Items	72
8.	Multivariate Analysis of Variance Illustrating Relationship Between Teaching Experience and Instrument Items	73
9.	Multivariate Analysis of Variance Illustrating Relationship Between Teacher Age and Instrument Items	75
10.	Multivariate Analysis of Variance Illustrating Relationship Between Education Level and Instrument Items	76
11.	Multivariate Analysis of Variance Illustrating Relationship Between Occupational Experience Other Than Teaching and Instrument Items.	77

Table		Page
12.	Purdue Rating Scale For Instruction	97
13.	Number of Student Ratings Collected for Pretest	99
14.	Number of Student Ratings Collected for Posttest.	100
15.	Gain Scores by Item for Each Group	101

CHAPTER I

THE PROBLEM

Description of the Problem Situation

Efforts to determine what constitutes effective teaching have not been productive. Currently there is no widely accepted definition of effective teaching nor a validated system for determining what constitutes an effective teacher (Cohen, 24:13). While research defining effective teaching continues, there appears to be little confidence among many professional educators that such research will be successful.

Despite this apparent lack of an acceptable definition for effective teaching, evaluation of teachers occurs. Educational administrators faced with problems of promotion, tenure, salary adjustment and dismissal are forced to arrive at judgments concerning effective teaching even though their judgments are subjective and based upon limited evidence and unvalidated procedures. Teachers, likewise, must evaluate themselves without the benefit of an acceptable definition for effective teaching.

Some evaluation efforts include instruments which identify desirable teaching behavior. Many of these efforts however, proceed without feedback which teachers receive to assist them modify undesirable teaching behavior. Such teacher evaluation efforts occur in many educational organizations and agencies in the United States.

The Virginia Community College System is one such agency faced with the problem of instructor evaluation. Each community college in the system is allowed to select its own evaluation process. Some community colleges utilize evaluation inputs from supervisors, peers, students or instructors themselves, while others depend more heavily upon supervisor evaluation. Although these evaluation inputs provide evaluation data to supervisors, they do not necessarily provide feedback to teachers for them to improve teaching performance.

Statement of the Problem

The problem of this study focused on utilization of feedback which enabled teachers to modify teaching behavior after receiving the feedback. Questions which were answered by the research were:

1. If teachers are provided behavioral feedback about their teaching performance and an opportunity to modify that behavior, will there be improvement in behavior measured by students during a subsequent performance?
2. What types of feedback currently used in community college settings have researchers reported?
3. Are specific types of feedback associated with behavioral changes?
4. Is a ten-week academic quarter sufficient to provide feedback and allow teachers time to make behavioral changes?

5. Can selected aspects of teaching behavior be identified and a change in these aspects of behavior measured?

6. Is there a relationship between selected demographic characteristics of the population and teaching behavior?

Objectives of the Study

Objectives formulated to answer the preceding questions are:

1. To provide behavioral feedback to instructors and statistically measure resultant behavioral changes.

2. To review current literature and report types of feedback used.

3. To select three types of feedback and statistically measure which types were associated with behavioral changes.

4. To choose a validated student rating instrument identifying and measuring aspects of teaching behavior.

5. To relate teaching discipline, teaching experience, age, highest degree earned and occupational experience with teaching behavior measured by students.

Assumptions

Assumptions made for this study were:

1. Teaching behavior could be measured quantitatively.

2. Instructors had a desire to improve teaching behavior.

3. Instructors made a self-analysis of received feedback to identify behavior needing modification.

4. Students were capable of assessing teaching characteristics identified by a validated student rating instrument.

5. Student responses on the rating instrument were accurate and unbiased.

Limitations of the Study

1. The population was limited to part-time and full-time instructors teaching on-campus credit courses at New River Community College, Spring quarter, 1975.

2. Students rating instructors were limited to those enrolled in the classes of instructors participating in the study.

3. Feedback was limited to student ratings, videotapings and a combination of student ratings and videotapings.

4. The evaluation instrument was limited to one which measures students' perceptions of the teacher and the classroom teaching situation.

5. Only those items on the evaluation instrument directly related to teaching behavior or teacher manipulation of the instructional environment were used,

Definition of Terms

Teaching behavior or performance - teaching behavior or manipulation of the instructional environment as identified by the Purdue Rating Scale for Instruction.

Modification - significant change in teaching behavior.

Teaching period or academic quarter - minimum of thirty hours teacher-student contact.

Closed circuit television (CCTV) - system including camera, videotape recorder and viewing monitor.

Pretests - student analysis of teaching behavior consisting of the instrument, Purdue Rating Scale for Instruction. Pretests were administered before feedback was provided.

Posttests - same as pretests except administered after feedback was provided.

Significance of the Study

Quantitative data supporting improvement of teaching performance may be needed if teacher evaluation is to be utilized objectively by college administrators and teacher supervisors. Without quantitative data about instructors' classroom behaviors it is often difficult for administrators and supervisors to make objective assessments of teachers' performances. A significance of this study was that it quantitatively measured changes in teaching performance.

Student ratings were reported to be used by many colleges throughout the United States to evaluate teachers' performances (Bryan: 14). It is questionable, however, if student ratings are useful measures of teachers' performances (Brickman: 16). This apparent reliance on student ratings as a means for measuring teacher behavior should not go untested. Another significance of this study was that it tested student ratings as a measure of changes in teachers' behaviors.

Many educators stated the primary purpose of evaluation is to provide teachers with information to improve their teaching performances (Borg: 14). The information received from evaluation serves as feedback which identifies teaching behavior. Feedback enables teachers to change undesirable teaching behavior. A significance of this study was that it compared three types of feedback and determined ones in which changes in teacher behavior occurred.

Another significance of this study was providing research about feedback in evaluation heretofore lacking in higher education (Pambookian, 55:52). Little research using feedback in junior or community college evaluation has been reported.

Microteaching successes with elementary and secondary teachers warranted research to determine its success on the college level (Waimon, 64:92). It was significant that this study incorporated microteaching as one of the feedback techniques.

A final significance of this research was its contribution to the body of professional knowledge concerned with teaching performance in the community college. Because the community college serves unique students, who for the most part have social, economic or academic deficiencies, teachers are expected to develop teaching strategies which serve unique students found at community colleges. Thus, it is essential that community college teachers receive accurate and quantitative feedback concerning teaching performance.

Organization of the Study

This study is organized into five chapters.

Chapter one contains a description of the problem situation, statement of the problem, objectives of the study, assumptions, limitations of the study, definition of terms, significance of the study and organization of the study.

Chapter two contains reviews of literature, summary and theoretical framework and statement of hypotheses.

Chapter three, Research Methodology, includes population and sample selection, research design, variables to be related, instrumentation and data collection, statistical analysis of data and a time schedule for administering the treatments.

Chapter four contains the results of the study.

Chapter five includes the summary, results, and recommendations.

CHAPTER II

REVIEW OF THE LITERATURE

A review of literature was undertaken to locate research regarding the use of feedback in faculty self-evaluation strategies and report what conclusions have been drawn about the effectiveness of those strategies. Literature was reviewed regarding:

1. The need for faculty self-evaluation.
2. The use of feedback in modifying teacher behavior.
3. Closed circuit television as a feedback source for faculty self-evaluation.
4. The combined use of closed circuit television with other feedback.
5. Student ratings as a source for evaluating teaching behavior.

Need for Self-Evaluation

Self-evaluation is one method by which an instructor can receive pertinent information he needs to improve teaching behavior.

Bayley (8) reported that colleague evaluation is a widely used form of teacher evaluation on the college level. Most teachers prefer peer criticism over student evaluation,

another widely accepted form of evaluation. Bayley (8) stated that colleagues are accepted as the most trustworthy judges of teaching performance. Despite acceptance of peer evaluation, concluded Bayley, most college instructors believe they are capable of criticizing their own teaching performance and many teachers prefer self-evaluation over other methods.

In an experiment utilizing microteaching, Naeslund (49) reported that teacher participants were critical of administrative problems which occurred. However, participants were extremely pleased with opportunities to evaluate their own teaching. Participating teachers stated that self-evaluation is needed for improvement of teaching performance.

Hrivnak (37) reported that a study testing a model of inservice education, using individual feedback to assist teachers in continuing their professional development, was successful in producing a positive change in teaching performance. The teachers' evaluations indicated their belief that the model was most useful in assisting them develop and refine teaching competencies.

Demands for quality education have illustrated the need for objective merit rating of teachers for promotion and salary increases. Bolin and Muir (13) reported a study which reviewed merit rating of teachers for promotion and salary increase. It was determined that a combination of

supervisor, peer, student and self-evaluation provides the best basis for decisions on merit increases and promotions. Twenty-one rating scales used in the study revealed nine factors involved in faculty ratings: classroom performance, personality traits, professional growth, student-faculty relationship, community service, service to the institution, length of service, research and publication and offers for employment from other institutions. There was no pattern of merit rating determined by the study. However, Bolin and Muir reported that there was agreement on the factors important in judging teacher effectiveness and these might serve as a guideline in developing merit pay schemes.

Dugas (26) reported that during an NDEA Institute where teaching participants viewed themselves on videotape it was determined that microteaching is an effective device in retraining experienced teachers. Although the participants taught only two fifteen minute class segments, they agreed that by being able to view themselves they could more fully appreciate constructive criticism offered by others, recognize their own strong and weak points and evaluate their own performance as teachers.

Redfern (57), in a paper presented at the 1969 annual meeting of the American Association of School Administrators, stated that teacher militancy will not necessarily be removed by placing emphasis upon higher salaries, better fringe benefits, a shorter work year, lighter teaching loads.

and other rewards. Rather, militancy may be reduced if ways are found to enable teachers to feel a sense of intrinsic worth and job fulfillment in the performance of their teaching duties. This, Redfern stated, would require a new approach to evaluation. He suggested that teachers and administrators agree upon specific, relevant performance objectives and that teachers have an opportunity to check periodically how well instructional procedures are achieving results. This would require some form of self-evaluation. Redfern concluded that the degree of militancy and the flexibility in evaluation are closely related.

In a guide intended to fill a gap in performance evaluation inservice training, Armstrong (5) stated that most evaluation materials have been prepared for the evaluators rather than for instructors, whose understanding of the evaluation process is vital to its success. The guide is written from the evaluatee's viewpoint and stresses the need for setting job targets, using monitored data, self-evaluation and evaluation conference preparation and participation. Armstrong also discussed the importance of behavioral objectives in setting goals, the power of self-evaluation, the requirements or documentation of performance data and the need for mutual trust between the evaluator and the evaluatee.

Bhaerman (10) presented reasons why most teachers will not accept evaluation for merit pay. The strong feelings of educators rejecting evaluation for extrinsic reward appear to be related to objective and quantitative measurements. Bhaerman contended that teachers will, however, support evaluation for constructive, diagnostic purposes only.

Young (68) stated that in modifying teaching behavior in the desired direction specific supervisory skills are required for various conference strategies. The provision of feedback is cited as being one of the most important aspects of the supervisor's strategy. Other strategies suggested are selection of focus, provision for discrimination training, provision of a model of teaching behavior and provision of closure. Videotape feedback is mentioned as one useful source of feedback since it provides a common frame of reference for both teacher and supervisor. Young held that the supervisory function should be one of "teaching teachers" rather than evaluation of teaching styles. The supervisor serves as a teaching consultant guiding the teacher through a systematic analysis and modification of teaching performance. This was accomplished via an analysis of self, learner, content and conditions of learning.

In the second of six components of the Instructional Staff Development Program, Urback (62) listed the main goals

of the program. They were introduction to interaction analysis, practice in microteaching and improvement of individual inquiry behavior. The program was designed to convince teachers that through self-analysis teachers can control their influence behavior in predetermined ways.

Fishman (28) stated that the evaluation of teaching effectiveness in the United States has concentrated on teacher characteristics and behavior. It is clear that teacher and method are important to the learning process but an effective teacher has not yet been defined. Teaching can not be evaluated independently of the individual teacher's behavior. Therefore, Fishman concluded, teachers must learn to accept self-evaluation as a means of teacher evaluation.

Harrington (33) reported that four feedback techniques were field tested on twenty-eight instructors in a technical institute. The four techniques, self-review, fellow instructor review, student review and teacher-educator review were determined as feasible for inservice programs. Also, teacher opinion of the techniques indicated the need for self-evaluation in teaching vocational and technical subjects.

Changes that have occurred in the area of teaching measurement and evaluation were discussed by Callahan (18). He suggested that developments in measurement and evaluation

in the past have become detached from the classroom situation. Although most evaluation schemes have continued to strive for objectivity and standardization they remain more concerned with group comparisons than with individual progress. Videotaping appears to have considerable potential as a tool in teacher evaluation. The need for self-evaluation in classroom related situations is becoming more widely recognized.

Ronan (58) reported that a device to measure college classroom teaching effective has been established using the critical incident technique. The device was derived from a study conducted at Georgia Institute of Technology in which three thousand incidents describing the "best" and "worst" teachers were collected. A taxonomy of teaching behaviors was isolated. These teaching behaviors can be used to describe effective and ineffective teaching as seen by students. The taxonomy was developed because of the need of instructors to involve themselves in self-evaluation.

Summary. The need for self-evaluation has been reported both from researchers advocating self-critique and from researchers describing utilization of self-evaluation techniques. It was reported that whether or not supervisors and administrators receive evaluation data about faculty does not greatly effect the manner in which faculty perform in the classroom. What counts is the degree and type of

feedback which the instructor receives about his or her teaching behavior. Only when teachers are provided opportunities to evaluate themselves can they change undesirable teaching behavior. It was also reported that extrinsic rewards, such as merit pay, promotion and tenure are not a motivating factor in bringing about a change in teaching performance. It was reported that such rewards may cause mistrust between teachers and supervisors. Self-improvement and personal growth were cited as primary reasons for the change of classroom behavior. It appears that evaluation systems which lack self-evaluation opportunities are not sufficient to bring about a change in teaching performance. Self-evaluation appears needed for teaching improvement.

Use of Feedback

Most of the literature reviewed stated that it is necessary for instructors to evaluate themselves. Feedback was reported as needed to enable teachers to make the self-assessment accurately.

There are several sources of feedback. The ones reported in the review of literature were student ratings, supervisor ratings and viewing one's self on videotape.

Larson and Ringness (42) reported a pilot project which evaluated aspects of teacher personalities by experimentally manipulating feedback. Ratings of teachers were given to the teachers and their supervisors to determine

whether such feedback included differential changes in the personality reassessment of the teachers. Personality assessment techniques were used with a group of eight elementary teachers. The subjects were divided into two treatment groups and one control group. Personality assessment was communicated to one treatment group but not to their supervisors. Assessment results of the group were given their supervisors at the same time as the instructors received them. The results were withheld from the control group until after the teaching sessions were completed. Evidence was found to suggest that the provision of personality information to the teachers and their supervisors was useful but no conclusions were drawn as to the effectiveness of the treatments to enable the teachers to make reassessments of their personalities.

Allen and McDonald (3) also failed to reach any conclusions regarding feedback in faculty self-evaluation. They report on a series of experiments designed to assess the usefulness of television recordings in improving reading performance. The experiments compare the effects of self-evaluation of teaching performance with feedback provided by a supervisor. No conclusions were drawn to support that feedback caused a greater change in faculty performance than the self-evaluation utilizing videotaping.

Some researchers claimed that feedback is not effective in causing a change in behavior. In a study to test the effect of feedback on behavior ratings, Hovenier (36) hypothesized that feedback would influence behavior. He tested his hypothesis by having two hundred and eight teachers rate their supervisors on ten factors considered desirable or undesirable for supervisors. Results showed that the total effect of feedback was less than expected.

Miller (46) concurred with Hovenier in a study which determined if providing instructors with information from student ratings affected their subsequent ratings by students. Thirty-six teaching assistants at the University of Iowa were used in this experiment. The results of the study indicated that student ratings did not improve performance. Miller concluded that a partial explanation of the findings might be the lack of career teaching commitment on the part of some teaching assistants, hence a lack of motivation to use feedback data to improve instruction.

Many investigators reported that feedback is an effective tool for self-evaluation.

Pachla (54) reported a 1968 study of instructors of vocational subjects in a post-secondary technical institute. The results demonstrated that those who received feedback showed greater gains in student ratings than did those who received no feedback. In another study reported by Pachla, a comparison in the change of student ratings from midtern

to the end of the semester for full-time college faculty was made to determine the effects of the feedback. Only behaviors which were expected to undergo changes in half a semester were rated. Results were that student feedback was effective in causing some change, although slight, in teacher behavior.

Another investigation on the effects of feedback from student ratings was reported by Centra (2). In this college-level study each of four hundred teachers from five colleges was randomly assigned to a feedback group, a no-feedback group or a posttest group. Teachers rated themselves and were rated by their students. The major hypothesis of the study was that student feedback would effect greater changes in teachers who rated themselves more favorably than did their students. Findings suggested that the greater the discrepancy between teacher rating and student rating where the teacher rating was higher than the student ratings, the greater the change in teaching behavior.

Gage and others (30) supported the conclusions of Centra that student feedback is effective. In a study using one hundred seventy-six Illinois elementary teachers who were divided into experimental and control groups, Gage found that when confronted with feedback about their behavior from students, and confronted with students' opinions

of an ideal teacher, instructors made significant gains in closing the gap between their students' concepts of them and the ideal teacher. An additional finding of the study was that the experiment enabled teachers to be more accurate in predicting pupils' concepts of the ideal teacher and thus were able to effect positive changes in their teaching behavior before the formalized feedback data were collected and disseminated.

Many teachers reported that they like feedback from students, mostly because it provides a non-threatening assessment about their teaching techniques and styles. However, if feedback becomes part of teachers' permanent records, or is submitted to supervisors for review, many teachers question the ability of students to accurately rate teachers.

Videotaping was reported as another source of feedback accepted as an effective tool in self-evaluation.

Perlberg and O'Bryant (56) reported that video recording provides authentic feedback of classroom interaction. Used in microteaching, videotaping provides a systematic method for analysis of teaching behavior. Perlberg and O'Bryant also reported that video feedback has resulted in improved teacher-classroom interaction at the college level. The development of individual and group models and their use motivated self-improvement in teaching strategy.

Although Perlberg and O'Bryant concluded that videotaping is an effective form of self-evaluation, they recommend teachers be aware of technical difficulties which might occur with videotaping. They suggest to encourage use of videotaping that tapes not be viewed by supervisors.

Hoehn (35) also documented the effective use of videotaping as feedback in self-evaluation. He described a program of inservice teacher self-improvement based on goal setting by the teacher. The program provides for the practicing of new skills, the reception of feedback regarding the effectiveness of the new skills and for self-analysis of the dissonance between goals and practice. Several field trials of the program have been conducted and results indicate that video feedback is effective in bringing about changes in teaching behavior.

Borg and others (14) concluded that feedback is effective in self-evaluation, and the degree of effect is proportionate to the amount of feedback received. He describes "Mini Course I," a short microteaching program designed to change twelve specific classroom behaviors. The program relies on videotapes illustrating model teachers and provides feedback through carefully structured teacher self-evaluation of the taped replays. The study was designed to estimate the degree to which the videotaped feedback influenced changes in teacher behavior. Teachers

were assigned to different treatment groups. Each group received a different amount of feedback, ranging from several feedback sessions to no feedback at all. It was concluded that those who completed all feedback sessions exhibited a greater change in behavior than those receiving less feedback. Those receiving no feedback exhibited the least amount of change.

Webb and others (66) concluded that feedback is so effective in self-evaluation that a program implementing large-scale microteaching at Brigham Young University will be expanded into a larger program. The success of a microteaching experience for seven hundred thirty student teachers based on eight to ten minute videotaped teaching sessions warranted the expansion of the program. Logistical problems in providing facilities and time for both taping and feedback sessions for over seven hundred students was reported the greatest obstacle to overcome in the program. In summary, it was reported that due to the success of feedback in self-evaluation sessions, the concern is no longer program justification but program expansion.

Summary. It was reported that feedback from at least one external source is necessary for faculty members to conduct an adequate self-assessment session of their teaching behavior. Two specific feedback sources have been identified as especially effective; student ratings and

videotaping. Videotaping was reported as the most valued by instructors since no observer bias can enter the ratings. However, videotaping was most feared as an evaluation source to be supplied to supervisors. Some debate about the effectiveness of student ratings was reported but most researchers agreed that videotaping and student ratings are effective for self-evaluation of teachers and can cause change in teaching behavior.

Use of Closed Circuit Television

Although many researchers agree that self-evaluation promotes change in teaching performance, they do not agree which type of feedback is the most effective for self-evaluation. One type reported as effective is microteaching. Microteaching provides teachers with the opportunity to critique themselves via videotape after they have presented a short mini-lesson of five to fifteen minutes duration. Other uses of closed circuit television less formalized than microteaching were reported as effective for self-evaluation.

Saloman and McDonald (60) investigated the effects of viewing videotapes of one's own teaching behavior. Thirty-eight teaching interns were permitted to view videotapes of their practice teaching sessions. Viewing questionnaires were completed by the interns before and after the viewing sessions to assess self-attitudes and attitudes towards teaching performance. The questionnaire results revealed

that satisfaction with one's own teaching performance was determined by initial self-attitudes. Although negative reactions were expressed on the questionnaires, the teachers' attitudes toward videotaping as a self-evaluation tool did improve. The findings of the study suggested that reaction to the tape was dependent on personality traits of the viewer.

Greenberg (32) stated that if ways to determine objectively the presence and effectiveness of desired teaching behaviors were developed, one could evaluate the instructional process with more validity and reliability. There are instruments which can be used to describe interaction in the classroom, but Greenberg doubted if these instruments can be translated into concepts which the evaluated teacher can visualize. Something other than written responses is necessary for teachers to evaluate themselves. He recommended the use of videotaping. Teachers can record their class sessions and use the tapes for analysis of teaching behaviors and relate the relationship of these behaviors to specific instructional objectives. The greatest value of videotaping is that the result is value free. The tape will visually and audibly depict what has occurred in the classroom session without evaluating that occurrence. Values are brought to the tape by the user. This means that neither student nor supervisor bias affects

the self-evaluation process. There is a danger, of course, that the teacher's own bias will cloud the self-evaluation process. Despite this danger Greenberg believed that videotaping is an effective means of self-evaluation.

Young (68) described microteaching as a scaled down teaching session lasting five to twenty minutes and utilizing one to five students. Although microteaching was first developed for preliminary experience and practice in teaching and as a research tool to explore training effects under controlled conditions, the method can serve the experienced teacher as a means of gaining new information about his teaching in a relatively short time. Videotaping the class session is an integral part of microteaching and immediately after the class session the teacher reviews his performance. He may make a self-evaluation, or a colleague or supervisor may participate in the evaluation. Videotape provides a common frame of reference for all evaluators so recall is not necessary for evaluation. Some instructors stated they would be uncomfortable viewing tapes with their colleagues or supervisors present. Young stated that the purpose of videotaping a microteaching session is to provide the teacher with a self-evaluation tool. Watching his own performance may suggest alternative approaches to strengthen teaching strategy. Videotaping is an effective tool for evaluation and should be available to all instructors.

Wagner (63) reported that microteaching is not an effective device for self-evaluation. He described a pilot project where teaching two sessions, reviewing the lessons on videotape and receiving feedback did not result in significant overall changes in teaching behavior.

Stukat (61) reviewed two Swedish microteaching experiments in which the first experiment sought to determine if microteaching was appropriate for inservice teacher training. The second instructed teachers to distinguish effective from less effective teaching behaviors. Flander's Interaction Analysis was utilized by teachers to rate themselves while viewing videotape. The investigators concluded that microteaching did not have lasting effects on teachers. However, teachers reported positive attitudes toward opportunities to evaluate themselves.

Langer (41) described a self-contained instructional package entitled, "The Mini Course," which was designed to train teachers to evaluate themselves in a microteaching situation. "The Mini Course" employed six videotaped microteaching sessions and was organized into four to six sequences during a three day program. The teacher was familiarized with specific teaching skills and was provided the opportunity to evaluate his performance of those skills. Each mini course was provided with a specific behavioral objective which was achieved by the accomplishment of a specific teaching skill.

Langer (41) stated two basic reasons which account for the success of "The Mini Course:" reinforcement derived from the emphasis on the perception of a behavior change and association with increased proficiency of a teaching skill.

Clayton and others (22) reported a two year study conducted at Syracuse University which attempted to ascertain whether summer workshop training of teachers in the use of feedback techniques such as Flanders Interaction Analysis, videotaping techniques and supervisory tools would produce a measurable change in the instructional behavior of student teachers. The results of the study showed an increase of congruence between the stated intent of the student teachers and their instructional performance over the period of student teaching. Clayton suggested that there was slight evidence that for definable subgroups of teachers particular activities were related to positive behavioral changes. The investigator concluded that positive changes in teaching behavior was brought about by the use of videotaping.

Cameron and Cotrell (19) reported that feedback techniques involving video recordings were tested with teachers serving isolated areas. Thirty-nine teachers from remote parts of Colorado were randomly selected and assigned to three treatment groups. The groups were video phone feedback, video mail feedback and video self-evaluation. A satisfaction scale and a reaction questionnaire were

administered. The conclusions drawn from the study were that the video phone or the video mail treatments were no more effective than the video self-evaluation.

Karasar (40) sought to test the impact of video feedback on teachers' eye-contact mannerisms. In micro-teaching sessions during an inservice vocational education workshop pre-recorded tapes from two groups were compared. One group had received video feedback, the other had not. The groups were compared on their introduction, presentation, summary of the teaching material and a combination of all three. It was concluded that video feedback may not have any significant effect on eye-contact mannerisms.

Dunay and Robison (27) reported that self-analyses of teaching behavior by elementary student teachers using the Baruch Behavior Analysis, were examined for accuracy, reliability, qualification and interpretation. The instrument was designed to help the student teacher analyze himself with the aid of videotape and prescribe improvement for his behavior. It was concluded that most student teachers recognized the need to improve their teaching behavior.

In an Associated Colleges of the Midwest pilot project (51) conducted at Carleton College, Northfield, Minnesota, one experiment was conducted to determine the feasibility of using portable, low-cost TV equipment to

record spontaneous learning-teaching activities and to experiment with videotaping in appraising the performance of student teachers. Various instructional techniques were utilized with both slow and gifted students. The teachers and their techniques were videotaped and reviewed by both teachers and students. Both teachers and students evaluated themselves and teaching techniques. It was concluded that student teachers profited greatly from self-evaluation permitted by viewing their own tapes.

Belt (9) reported that microteaching at Brigham Young University consisted of a lesson presented by a student teacher to a microclass of three to five high school students. In some cases peer teachers were utilized in the microclass. The members of the microclass evaluated the teacher's performance. The performance was videotaped so that the teacher could evaluate himself. A reteach session was conducted several weeks later using different students. The reteach session was also videotaped and evaluated. Conclusions drawn were that provision for immediate feedback and self-evaluation were highly important and that teacher performance was improved following evaluation and playback of the videotape.

Kallenbach (39) reviewed a research study of a microteaching method developed at Stanford University. The method consisted of presenting four to six minute videotaped lesson segments which were viewed and discussed by

the instructors and their supervisors. Each lesson was immediately retaught with a comparable group of students. A critical analysis followed this second videotaped session. It was determined that the feedback obtained from the videotaped session was relatively ineffective as compared to the feedback provided by the supervisors.

Moe, Alden and Feehan (47) reported on the use of videotape recorders in the developmental reading practicum fellowship program at Clarke College, Dubuque, Iowa. Through microteaching the participating teachers developed technical teaching skills necessary in content areas of the developmental program. To achieve this, a videotape was made of the teacher as he taught both small and regular classes of students. The videotape was viewed, discussed, and evaluated by the teacher and the supervisor. After several sessions the teacher was able to improve teaching methods.

Johnson (38) discussed an innovative inservice training program at Wheeling High School, Wheeling, Illinois, emphasizing the affective domain of teaching techniques. The main objective of the program was to expose teachers to situations which would result in their being willing to evaluate their own teaching behavior. The classroom sessions were videotaped and participating teachers who were given opportunities to view their videotapes were willing to evaluate their teaching behavior.

Codwell (23) reported on a project designed to demonstrate the effect of adopting microteaching techniques on the instructional behavior of rural school teachers in Tennessee, Georgia, and Florida. The teachers were pre-rated by observers using the Stanford Teacher Competence Appraisal Guide. Each teacher evaluated his own attitude using the Minnesota Teacher Attitude Inventory. Teachers' performances were recorded on videotape which were then viewed by the teachers. It was concluded that a rural school teacher's opportunity to see his performance by means of videotape resulted in an improvement of his instructional behavior.

Gall and others (31) reported a study which determined whether personality variables are correlated with specific teaching behaviors when measured before and after a microteaching mini course. Teaching behavior was measured while each teacher taught a twenty minute videotaped lesson in his regular classroom. The videotapes were scored by trained raters who rated personality variables of the teachers. The implications of the study were that personality has little effect on teaching behavior of females but a strong effect on males. Also, the study implied that microteaching utilizing videotaping may reduce individual differences in style, thereby partially canceling out the effect of personality.

Roush (59) reported that videotaping as a feedback device was not able to produce behavioral changes in a group of teachers significantly different from those who did not receive feedback.

Self-confrontation by means of videotape recording has become a popular technique for allowing student teachers to discover ways in which their own behavior influences the teaching process. Bierschenk (11), reported ways in which the student teacher perceives and evaluates his own behavior. The results of the study revealed no differences in the objectivity of a student teacher's self-assessment whether he used self-confrontation via videotaping or not.

Friebel and Kallenbach (29) reported that teaching behavior can be changed with feedback. As a result of participation in a special mini course it was hypothesized that there would be no difference between student teachers receiving the entire special mini course and those omitting the videotaping session from the mini course. The hypothesis was rejected and it was concluded that videotape feedback is necessary for successful completion of the special mini course.

Young and Young (69) discussed the effectiveness of individually prescribed microteaching training modules. The study, conducted at John Hopkins University, compared

pre-internship student teachers with those serving their internship. Each group was divided into those who involved themselves in a microteaching session and those who did not. It was concluded that those who participated in a microteaching session in both groups acquired a significantly greater number of behavior skills.

Baird (6) reported a study in which each student teacher was videotaped during a teaching session, and then attended an evaluation session where detailed and critical comments were presented by a supervisor and the class. It was concluded that combining videotaping and class critique, in addition to the self-evaluation, was more effective than videotaping alone.

Summary. Closed circuit television has been reported as an effective tool for self-evaluation, and its use in formal microteaching was frequently reported. Less formal use of videotaping has been reported in other self-evaluation situations. Not all investigators reported significant changes in teacher behavior caused by videotaping but many agree that the use of closed circuit television is accepted by teachers as helpful in changing teaching performance.

Some researchers reported evaluation procedures utilizing only videotaping while others reported combining its use with peer critiques or supervisory evaluation. In

these reports, however, videotaping was the common element which enabled the evaluation session to be successful.

The reported usages of videotaping were on elementary, secondary and college levels. It was reported that closed circuit television is an effective device for self-evaluation in each of these levels.

A Combination of Closed Circuit Television and Other Feedback Devices

Although closed circuit television has been reported effective as a self-evaluation tool, reports indicated that its combination with other feedback sources results in a more significant change of instructor performance. Other feedback devices reported were supervisor or student ratings, peer or supervisor critique sessions and interaction analysis by students, colleagues, or supervisors.

Oliver (53) reported a study which does not include closed circuit television but does support the use of combined feedback techniques. The study determined the effectiveness of informational feedback from supervisors, students and a combination of supervisors and students as a means of improving teacher performance. Two hundred eighty-six vocational teachers in New Jersey were categorized into groups based on years of teaching experience and were then randomly assigned to one of four groups receiving different feedback conditions. The types of feedback were from supervisors only, senior high school students

only, a combination of supervisors and students and a control group where no feedback was supplied. Findings supported that the student-only group and the combined group supplied feedback which caused a significant positive change in teacher effectiveness.

Oliver's conclusions were supported by Aleamoni (2) who conducted a study to assess effects of a combination of student feedback and faculty consultations. The student feedback was collected and consultations were conducted during one semester of college instruction. The conclusion drawn was that by providing student feedback along with faculty consultations the instructors could significantly improve their student ratings.

Ward (65) compared the difference in teaching behavior among three groups of teachers who utilized different feedback techniques. Inservice elementary teachers were assigned to receive feedback either from videotapes, reflective evaluation, or a combination of the two. The findings were inconclusive, but Ward stated that teachers expressed greater satisfaction with the combined mode than with either of the other two.

Barney (7) agreed with Ward in that the combination of closed circuit television and other feedback techniques in evaluation is more effective than one technique used independently. To increase the efficiency, relevancy, and effectiveness of preservice and inservice training in home

economics education, a one-year project incorporated videotape feedback into the already established evaluation process. Videotaping was incorporated in three of five programs. Student teachers in the three programs received a one semester microteaching experience in addition to the other evaluative techniques, including supervisory and pupil evaluation. It was concluded that the inclusion of closed circuit television in the self-evaluation program was effective in improving the student teacher's behavior.

Meers, (45) likewise, supported the use of a combination of videotaping and other feedback tools in teacher evaluation. He reported a study which determined if a difference existed in the level of psychomotor achievement among junior high school students receiving three different kinds of feedback. The treatments included viewing one's own performance on videotape, receiving critique from a panel of judges while viewing one's self on videotape and viewing the videotape with critique. It was found that students who viewed the videotapes plus received the panel critique had significantly higher levels of psychomotor performance in the final analysis.

Amidon and Rosenshine (4) reported on a model used for inservice teacher training and, like Meers, concluded that a combination of closed circuit television and other techniques promote maximum effectiveness in teaching techniques.

The model included both microteaching and interaction analysis and emphasized various teaching behaviors to be practiced. Student preparation for activities, acceptance of student feeling and acceptance of student ideas were skills evaluated. Amidon and Rosenshine concluded that the model was successful in changing teacher attitudes and behaviors, and that the combination of interaction analysis and videotaping produced an effective model.

McDonald (44) utilized videotape feedback in an experiment designed to determine the optimum combination of modeling and video feedback. McDonald assumed that video feedback is effective both as a self-evaluative tool and as a modeling tool. His experiment provided teachers with a video taped teaching model. A teacher presentation was videotaped and then the teacher compared his own session with the model. McDonald concluded that whenever the modeling cues were highly discriminative the feedback from the teacher's own videotape was most effective.

A study by Morse and others (48) utilized a combination of feedback techniques to change teaching behavior. Audio taped sessions were used in combination with supervisory feedback in the development of beginning teachers' refocusing behaviors. Although audio taping lacks the visual dimension of videotaping it was reported to be successful when used in conjunction with supervisory feedback. Four groups of teachers were formed and each

received a unique type of feedback. One group listened to their audio tapes but received no other feedback and another group received no feedback at all. The third group listened to their tapes and used a listening guide designed to aid the teacher pinpoint behaviors and the last group listened to their tapes, used the listening guides and received supervisory critique. The results of the study revealed that teachers who used audio tapes alone or with the listening guide exhibited no change in teaching behavior. Teachers in the group which received all types of feedback exhibited changes in their behavior.

Summary. It has been reported that a combination of feedback techniques used to evaluate teaching behavior is more effective than one method of feedback used alone. The use of closed circuit television in combination with one or more other self-evaluation devices has been reported as an effective format for altering the behavior patterns of teachers. Microteaching, utilizing short teaching sessions, was the most frequently reported use of closed circuit television combined with other feedback techniques. Other uses of closed circuit television coupled with supervisory and student evaluations were also reported as modes for stimulating behavioral changes in instructor performance. It was reported that a combination of closed circuit television and other feedback techniques is more effective in changing teacher behavior than is one feedback technique used alone.

Student Feedback

Student feedback, mostly in the form of student ratings, was reported as the most widely used type of feedback provided to teachers about their teaching behavior. Although student ratings are widely used they do not go unchallenged as a reliable method for teacher evaluation.

Abramson (1) reported that three of every ten colleges responding to a 1966 College Management survey stated they presently use a formal system of faculty ratings by students or would institute one in the near future. Some institutions reported using all grade levels of students. Others reported using only seniors or honor students. No attempt was made in the survey to ascertain the accuracy of student ratings.

Bryan (17) concurred with Abramson that many colleges are using student ratings. In a 1966-67 survey, in which three hundred and seven institutions responded, forty-nine percent of the responding colleges utilized an evaluation plan incorporating student appraisal of teaching. Only thirteen percent of the colleges reported that they had discontinued student ratings after a trial period.

Some researchers have delved into the question of student rating reliability. Lauroesch and others (43) reported a study which evaluated formats for providing

feedback of student opinion during teacher training. Sixty-nine teacher interns were randomly assigned to five treatment groups. One group received no feedback, one written feedback only, another, written feedback plus a conference with the supervisor, a fourth received delay treatment not administered until near the end of the semester and the last group received treatment at the termination of the semester. The feedback consisted of student responses to the Bryan Student Opinion Questionnaire. The purpose of the study was to detect possible sources of invalidity, and to assess the accuracy of the feedback. The results of the study revealed inconsistent student assessments and Lauroesch concluded that student assessments were not reliable.

Brickman (16) also questioned the reliability of student evaluation of faculty performance. He reported a rapid increase of student evaluation in colleges to rate college teachers knowledge of subject matter, teaching effectiveness and personality. Brickman questioned if students have the knowledge, experience and ability to evaluate college instructors with a proper degree of accuracy and fairness. He advised faculty and administrators to devise careful, periodic and objective evaluation schemes utilizing, but not depending entirely upon, student ratings.

Borgatta (15) noted that student rating procedures may be useful in helping to learn if a teacher is negligent

in the performance of his teaching duties but the ratings should not be used as evidence in making judgements about comparative effectiveness, promotion, tenure and salaries. Borgatta stated that colleges are institutions of higher learning and therefore emphasis should be placed on what the student learns and not what instructors teach.

Costin and others (25) concurred with Borgatta and Brickman in their review of findings concerned with reliability, validity and usefulness of student ratings. The review substantiated that student ratings can provide reliable and valid information on the quality of courses and instruction. Such information can be of use to academic departments to aid the individual instructor improve his teaching performance. However, student ratings fall short of a complete assessment of an instructor's teaching contribution.

Not all researchers agree that student feedback is biased or inaccurate. Many researchers accepted the validity and reliability of student assessments of faculty performance.

Novak (52) stated that students give accurate feedback of teaching performance. In three separate, sequential three-month studies, audiotapes were made of the teaching performance of eighteen teachers before, during and after teachers introduced feedback instruments to their students. The tapes were analyzed for changes in teacher behavior over the period of time the feedback was used. It

was concluded that students can give accurate feedback which effects change in teacher behavior, and that this change can be quantified and described.

Hassee and Miller (34) concluded that student ratings are a reliable source of feedback for faculty evaluation. In a study which evaluated both the view that one's knowledge of one's subject is enough to make an effective teacher and the view that teaching is an art rather than a skill, both views were challenged when results showed that teachers need continuous interaction with students to be effective. Teachers with access to student appraisal improved more significantly than those without the student input. In fact, student ratings were judged as the most stable and reliable measure of teacher behavior, even more so than supervisor ratings.

Nowell (50) agreed that students are good raters of teachers' performances. He reviewed findings on the effectiveness of student evaluation. These findings were based on studies which investigated how ratings were obtained, how qualified students were to rate teachers and how accurate the students were in their responses. He concluded that despite variables, such as class size and whether the class is elective or required, students are fairly good raters of their teachers.

Wise (67) discussed two viewpoints; one which expressed the belief that students cannot distinguish effective teaching from poor teaching, and the other which espoused

the tenant that students' reports of teachers are biased by such factors as grade expectations, student leniency and popularity of particular teachers. The first viewpoint is dismissed by Wise as totally unsubstantiated. He argues that the student has little opportunity to make value judgments of a teacher's effectiveness when standardized rating instruments are utilized. The student is merely a data generator in which he selects a level for a predetermined criterion and marks a response on the data sheet. The student does not participate in the selection of the criteria for the rating scale and in many cases is unaware of the meaning which is ascribed by others to his marks on a particular scale.

About the second view point, Wise stated that bias may be present in students' ratings. Wise cited several studies which resulted in a strong positive correlation between students' grades and students' ratings of teacher performance in which "A" students gave the highest ratings and "F" students gave the lowest. Wise conceded, however, that other studies have found a less positive relationship.

Summary. It was reported that educators do not agree which techniques are an accurate source of information for use in making decisions about tenure, promotion and pay. Many citations did not distinguish between reliability and accuracy as separate characteristics of student assessment of faculty performance. Those that did make a distinction

implied that although students are accurate in responding to rating questionnaires, they are not necessarily reliable.

Summary and Theoretical Framework

Summary. The literature reviewed illustrated the lack of agreement among researchers about the effectiveness of student or videotape generated feedback in faculty self-evaluation. However, there is sufficient agreement on aspects of self-evaluation to draw several conclusions.

The need for formalized self-evaluation by which an instructor can receive pertinent information to improve teaching performance was reported. Extrinsic rewards such as merit pay, promotion and tenure, were reported as not being motivating factors in bringing about improvements in teaching performance.

The use of feedback from sources other than the instructor himself was reported as necessary to provide accurate information about teaching performance to an instructor.

Videotaping was reported by researchers as being one of the greatest promoters of effecting change in teaching performance. Microteaching, a formalized scheme for providing teacher training sessions, provides immediate feedback of teachers' performances, and is relevant and unbiased. Many researchers agree that closed circuit television is accepted by teachers as helpful in changing teaching performance. Reported usages of videotaping included elementary, secondary and college levels.

A combination of several feedback sources was reported as more effective in bringing about changes in teachers' performances than any one single feedback device. Closed circuit television in combination with student ratings, supervisor ratings or peer critique sessions was reported as the most effective.

Student feedback in the form of student ratings was reported as the most widely used source of feedback provided to teachers about their performances in the classroom.

Conclusions drawn from the review of literature are:

1) Self-evaluation is necessary if teachers are to be expected to modify undesirable teaching behavior, 2) a source of feedback other than self-evaluation is required for an accurate and unbiased assessment of teaching performance, 3) closed circuit television provides an excellent source of feedback for self-evaluation, 4) a combination of videotaping and other feedback sources is a more effective feedback source than one source alone and 5) student ratings are widely used sources of feedback.

Theoretical framework. Based on conclusions drawn from the review of literature it is theorized that:

1. Immediate feedback is related to behavioral changes in instructor performance.

2. An improvement in teaching behavior will occur after feedback is provided.

3. A greater change in teaching behavior will occur

when videotape feedback is provided than when student rating feedback is provided.

4. A greater change in teaching behavior will occur when a combination of videotape and student rating feedback is provided than when either of the two is provided alone.

Theoretical schema. Below is an illustration of the theoretical framework:

INPUT	THROUGHPUT	OUTPUT
Immediate feedback about teaching behavior.	→ Given to teacher	→ A change in teaching behavior.
Initial instructor performance.	→ Feedback given teachers.	→ Improvement in teaching behavior during subsequent performance.
Videotape treatment vs. student rating treatment.	→ Quantitatively measure each as a change agent.	→ Videotape more significant as change agent than student ratings.
Combination of student ratings and videotaping vs. either of the two used alone.	→ Quantitatively measure each as a change agent.	→ Combination is more significant as change agent than either used alone.

Statement of Hypotheses

Major hypotheses formulated from the theoretical framework are stated in null form:

1. No significant difference in teaching behavior measured by student ratings shall occur between student

ratings administered before and those administered after feedback is given the teacher about his teaching performance.

2. No significant difference in teaching behavior measured by student ratings shall occur among teachers given different types of feedback.

CHAPTER III

RESEARCH METHODOLOGY

Introduction

Research methodology used to meet objectives of the study was presented in this chapter. The objectives were researched by testing two null hypotheses: 1) No significant difference in teaching behavior measured by student ratings shall occur between student ratings administered before and those administered after feedback is given the teacher about his teaching performance and 2) No significant difference in teaching behavior measured by student ratings shall occur among teachers given different types of feedback.

A description of the population and sampling techniques was presented. The research design was described and threats to internal and external validity were discussed. Treatments or independent variables, instrumentation and data collection were described. A time schedule for administering treatments and collecting data was detailed.

Population

This investigation was conducted during the Spring quarter of the academic year, 1974-75 at New River Community College, Dublin, Virginia. New River Community College is one of twenty-three community colleges in the Virginia Community

College System and serves the counties of Pulaski, Giles, Floyd and Montgomery and the City of Radford. New River Community College is a comprehensive community college with programs for college parallel, vocational-technical, business and secretarial and adult students.

The population of this study consisted of seventy-six full and part-time instructors who taught on-campus, credit courses and were employed during the 1975 Spring quarter. The population included teachers in all programs.

Research Design

The research design for this study tested each of the hypotheses and minimized threats to internal and external validity.

The design was a modification of the Solomon Four-Group which consisted of comparing pretests with posttests from five randomly selected groups. This design incorporated controls for internal and external validity. These controls are described under Controls for Threats to Internal Validity. The advantage of this design was that by paralleling experimental groups with a control group lacking the pretest, both the main effects of testing and the interaction of testing and the treatments are determinable. Generalizability was increased with this design. The modification made to the Solomon Four-Group was the addition of a pretest to the second control group. Whereas the second control group without a

pretest in the Solomon Four-Group design was to isolate the interactive effect of pretesting and the treatment, the modification was designed not only to isolate the interaction effect but to allow measurement of the effect.

A diagram of the research design is illustrated by Table 1.

Treatments. The independent variables in this study were three types of feedback manipulated in treatments given to the groups. Treatments administered were:

1. Treatment #1 (X_a). Treatment number one consisted of feeding back each instructor's pretest scores from the Purdue Rating Scale for Instruction to the instructor so that he had an opportunity to modify any undesirable behavior as indicated by those scores. The scores were the mean scores of all student ratings for each item of the instrument.

2. Treatment #2 (X_b). Treatment number two consisted of videotaping one of each instructor's class sessions for one hour and allowing each instructor the opportunity to view the tape once. Instructors did not have access to the pretest scores.

The videotaping equipment consisted of one video recorder and one camera. The camera was stationed in the rear of the classroom, with the lens adjusted to include the widest angle possible. No camera operators were employed.

TABLE 1

Diagram of the Research Design

Group	Sample Selection	Pretest	Treatment	Posttest
Control Group	(r)	0 ₁		0 ₂
Treatment #1	(r)	0 ₃	X _a	0 ₄
Treatment #2	(r)	0 ₅	X _b	0 ₆
Treatment #3	(r)	0 ₇	X _c	0 ₈
Second Control Group	(r)	0 ₉		0 ₁₀
Posttest Only	(r)			0 ₁₁

Legend. The symbols utilized are:

(r) illustrates that a random selection technique was used to assign subjects to the groups and assign the groups to the treatments.

0₁0₃0₅0₇0₉ represent the pretest which was administered. The pretest was the Purdue Rating Scale for Instruction:

X_aX_bX_c denote treatments which were manipulated. The treatments are described under the heading, Treatments.

0₂0₄0₆0₈0₁₀0₁₁ represent the posttest which was administered. The posttest was the Purdue Rating Scale for Instruction.

The video recorder was switched on at the beginning of the class session and switched off one hour later. No guide was given instructors as to the type of behavior they exhibited.

3. Treatment #3 (X_c). Treatment number three consisted of a combination of both treatments #1 and #2. The two treatments were given instructors in such a way that they had access to their student ratings and video taped performances simultaneously.

Controls for Threats to Internal Validity

The research design incorporated controls to minimize threats to internal validity. The threats are discussed as follows:

1. History. Any extraneous variables resulting from unwanted specific events occurring during the experiment were determined by a significant difference between O_1 and O_2 in the control group.

2. Maturation. Any processes which occurred within the respondents, operating as a function of the passage of time, were also determined by a significant difference between O_1 and O_2 in the control group.

3. Testing. The effects of pretests upon scores of the posttests were minimized by any significant differences occurring between posttest scores of the control group which received a pretest but no treatment and posttest scores of the posttest-only group which received no pretest nor treatment.

4. Instrumentation. The use of a standardized objective test, the Purdue Rating Scale for Instruction, and precise administration of the instrument precluded any possible errors which resulted in a significant difference between pretest and posttest scores.

5. Statistical regression. Random stratified selection of the sample reduced the possibility of selection of respondents whose extremes might cause regression of scores toward the mean.

6. Experimental mortality. The pretest scores of instructors unable to complete the experiment were retained. The mean of posttest scores of other instructors in the same experimental group was used as a posttest score for the mortally effected instructor.

Pretest scores of students unable to complete the experiment in each group were also retained. An average of the other students' posttest scores were substituted for mortally affected students.

7. Selection bias. Equal-size stratified sampling of instructors for each experimental group was conducted. Students for each group were those students assigned by the institution to each particular teacher. No random selection of students was possible.

Controls for Threats to External Validity

The research design incorporated controls to minimize threats to external validity. The threats are discussed

as follows:

1. Population invalidity. This study was limited to the population of New River Community College during the Spring quarter, 1975. Generalizations were made to that population with confidence. Generalizations might also be made to other institutions where it can be determined that the student and teacher population are similar.

2. Interaction of demographic variables and treatment. Equal size stratified random sampling was employed to negate possibilities that demographic variables of those receiving treatments would effect the study. The demographic variables considered in this study are discussed under the heading, Sample Selection.

3. Ecological invalidity. Ecological validity, or the degree to which generalizations can be made in another setting, was controlled by quantifying independent and dependent variables and by the precise description of the implementation of all aspects of this study.

Sample Selection

Six instructors were drawn from the population for each of six groups using an equal-size stratified random sampling technique. This technique served to negate possibilities of demographic characteristics affecting results due to small numbers of instructors in each treatment group. The following demographic characteristics were identified as being significant in affecting teaching behavior:

1. Teaching Area. Two teaching disciplines, vocational/technical and college transfer, were identified as characteristics. Approximately half of the population fell into each group.

2. Length of Teaching Experience. Previous teaching experiences on the high school or college level were considered characteristics.

3. Age. The ages of instructors were considered characteristics.

4. Education Level. The education levels teachers attained were considered characteristics.

5. Occupational experience. Occupational experiences other than teaching were considered characteristics.

Any treatment group dominated by instructors who possessed the same demographic characteristics, might be considered affected by the treatment and/or the demographic characteristics rather than by the treatment alone.

Characteristics, (1) Teaching Area and (2) Length of Teaching Experience, were considered as major characteristics and would most likely affect results of the study if they were not equally dispersed in the treatment groups. Factors, (3) Age, (4) Education Level, and (5) Occupational Experience, were considered as minor characteristics and would affect results only if one or more characteristics dominated treatment groups. Therefore, sample groups were stratified according to major characteristics. Minor characteristics were allowed to fall in the various groups by chance.

The population was divided into two groups according to teaching areas, (a) Vocational/Technical and (b) College Transfer. Each of these two groups was subdivided into three categories: (a) teaching experience less than two years, (b) Teaching experience from two years to four years and (c) teaching experience more than four years. Table 2 illustrates the number of population subjects according to teaching area which fall into each of the three subdivisions of teaching experience.

Six subjects were randomly chosen from each of the categories previously described and assigned to the treatment and control groups. Table 3 illustrates the results of sample selection denoting the number of sample subjects falling into each stratum of the five demographic categories.

Instrumentation and Data Collection

Instrumentation. The dependent variable in this study was teaching behavior identified by the gain scores of nineteen items on the instrument, Purdue Rating Scale for Instruction. Gain scores consisted of the difference between pretest scores and posttest scores. The gain scores of the Purdue Rating Scale were dependent upon the treatments administered. The Purdue Rating Scale was the only standardized instrument utilized in the study.

Purpose of the Instrument. The purpose of the Purdue Rating Scale is to measure students' perceptions of the teacher and the classroom teaching situation. The scale was

TABLE 2

Number of Population Subjects According to
Teaching Area Which Fall Into Each of the Three
Subdivisions of Teaching Experience

Teaching Area	Teaching Experience		
	2 Years	2-4 Years	4 Years
Vocational/Technical Occupational	10	12	11
College Parallel	13	14	16

TABLE 3

Number of Sample Subjects for Treatment and Control Groups
Falling into each Stratum of Five Demographic Characteristics

	Teaching Area		Teaching Experience			Age			Educational Level			Occupational Experience Than Teaching		
	College Parallel	Vocational Occupational	2 Years	2-4 Years	< 4 Years	> 31 Years	31-40 Years	< 40 Years	> BA Degree Plus 15 Quarter Hours	BA Degree Plus 15 Qtr. Hrs. MA Degree Plus 15 Qtr. Hrs.	< MA Degree Plus 30 Quarter Hours	> 4 Years	4-6 Years	< 6 Years
GROUP #1 Combination Feedback	3	3	2	2	2	2	2	2	1	4	1	2	2	2
GROUP #2 Student Rating Feedback	3	3	2	2	2	4	1	1	1	5	0	4	1	1
GROUP #3 Control	3	3	2	2	2	3	1	2	0	3	3	2	1	3
GROUP #4 Video Feedback	3	3	2	2	2	2	3	1	0	5	1	3	0	3
GROUP #5 Second Control	3	3	2	2	2	3	3	0	2	3	1	4	2	0
GROUP #6 Posttest Only	3	3	2	2	2	4	0	2	0	5	1	1	2	3

developed to give instructors a basis for self-improvement via feedback.

Reliability of the Instrument. The reliability of the Purdue Rating Scale is a function of the number of students doing the ratings. For ten instructors' traits with a class of twenty students, reliabilities (N=205 teachers) range from .84 (fairness in grading) to .95 (interest in subject). The overall rating by twenty students has a reliability of .91.

Validity of the Instrument. There is no available outside criterion through which to validate the Purdue Rating Scale. Since the students' attitudes and perceptions are at issue, validity by definition can be equated with reliability.

Item omission. Several items of the Purdue Rating Scale were omitted from the student ratings. Those items not pertaining to teaching behavior and those items over which the teacher had no control were eliminated.

The items eliminated were:

1. Item number twelve. "Suitability of the size of the class (consider the subject matter and type of class lecture, lab, etc.)." Class size is an administrative determination and was not controlled by the instructor.

2. Item number sixteen. "Suitability of the laboratory facilities available for the course." The laboratory facilities were limited and suitability was not controlled by the instructor.

3. Item number seventeen. "Suitability of the assigned textbook." In many cases the instructor had no input into the choice of the textbook.

4. Item number twenty. "How the course is fulfilling your needs (consider your ultimate as well as your immediate goals)." The instructor had no control over goals of the students.

5. Item number twenty-one. "Range of ability in the class (are there too many extremely dull or extremely bright students)." Student assignment was not controlled by the instructor.

6. Item number twenty-three. "The weight given to tests in determining the final grade for the course." The instructor determined grade weights and did not necessarily reveal weighing techniques to students.

Item omission did not affect instrument validity and reliability since the validity and reliability of the instrument is a function of the number of students doing the rating, not the number of items.

Data Collection. Students marked their responses to the instrument on optical scan sheets. Data on the sheets were electronically transferred to computer punch cards. The computer cards were used to compute the statistical analysis of data. Computation of this type data at the Virginia Polytechnic Institute and State University computer center required that student responses

be limited to no more than five choices per response. Therefore, responses to the instrument, the Purdue Rating Scale for Instruction, were limited to extremely poor, below average, average, above average and excellent. This limitation did not alter the validity or reliability of the Purdue Rating Scale since the data collected for this study were not compared with national norms.

Statistical Analysis

Treatment of Data. Ratio data were collected from both pretests and posttests. Each of the twenty items of the Purdue Rating Scale for Instruction was assigned five value levels from one point, being the lowest or poorest value, to five points, being the highest or best value. All data were computed on the basis of one point intervals between levels with zero representing no change from pretest to posttest scores. Changes from pretest to posttest scores were computed as gain scores. Gain scores served as the criteria upon which the statistical analyses were conducted.

Computation of Gain Scores. Gain scores were computed for each item of each group. Gain scores reflected either a positive or negative value depending on whether instructors' scores increased or decreased from pretest to posttest. Gain scores consisted of the mean of all students' posttest scores for each item minus the mean of all students' posttest scores for each item. Gain scores were computed for all treatment/

control groups except the posttest-only group. Gain scores for item twenty-six were not computed. It was determined as the study progressed that item twenty-six, "the overall rating of the instructor," would not yield specific information about teaching behavior. Its continued inclusion in the study did not justify the additional degree of freedom that would be lost in the statistical computation.

Test of the Hypotheses. The first hypothesis, no significant difference in teaching behavior as measured by student ratings shall occur between student ratings administered before and those administered after feedback is given the teacher about his teaching performance, was tested at the .05 alpha level with a multivariate T^2 test pairing the pretests and posttests of each of the nineteen instrument items. The program used was that described by Clyde Y. Kramer, "4.10 Multivariate Test of $H_0: \underline{0} = \underline{0}$ when Observations are Paired," A First Course in Methods of Multivariate Analysis, Virginia Polytechnic Institute, 1972, P. 96.

To test the second hypothesis, no significant difference in teacher behavior measured by student ratings shall occur among teachers given different types of feedback, at the .05 alpha level, three statistical steps were needed. Step one determined if a difference occurred among any of the posttest scores of the treatment groups taken together.

This step employed a one-way multivariate analysis of variance using the nineteen gain scores as the dependent

variable. The program used was the Statistical Analysis System, "Manova," by Anthony James Barr and James Howard Goodnight, Department of Statistics, North Carolina State University, Raleigh, North Carolina, August, 1972.

Step two determined in which treatment groups the difference occurred.

This step employed linear contrasts comparing the groups by twos. The program used was that described by Kramer, "5.10-Linear Contrasts Among Vectors of Treatment Means." The formulas used were $H = (N_1 \ C_1^2)$ Where N is the number in each group and C is the contrast value, and $U = \frac{(E)}{(E+H)}$

Where U is Wilks' Criterion, E is the error matrix and H is the hypothesis matrix.

Step three determined in which of the nineteen instrument items a difference occurred. This step incorporated simultaneous confidence intervals for differences between treatment means for each of the nineteen variables or items. The program used was that described by Kramer, "5.0 - Simultaneous confidence intervals for differences between treatment means," P. 139.

Relationship of Demographic Characteristics

Nominal data were collected for each of the five demographic categories identified for this study. Each category was stratified and a nominal value was assigned to each stratum for computational purposes.

A one-way multivariate analysis of variance was conducted for each demographic category identified in the study.

The Statistical Analysis System program computed an overall MONAVA for each of the demographic categories and an "F" approximation of the relationship that the demographic characteristics had on each of the nineteen instrument items. The approximation was computed as an "F" value of the independent and dependent variables but excluded interaction between the two. The level to which this approximation can be accepted with confidence was computed by the formula:

$$1 - (1 - \alpha)^P$$

Where α is the .05 alpha level established for this study and P is the number of dependent variables. The computation of this formula is as follows: $1 - (1 - .05)^{19} = .6226$. Therefore, the approximated "F" values for each of the demographic characteristics were computed at a .6226 confidence level and were not acceptable at the .95 confidence level stipulated for this study.

Time Schedule

The following time schedule was utilized:

Phase I (April 2-20) Identifying Population and Sample Selection.

- Step 1 Identify all full-time and part-time teaching faculty teaching on-campus courses.
- Step 2 Randomly select thirty-six faculty for six groups.
- Step 3 Randomly assign each group to a treatment.
- Step 4 Determine the number of students taught by each instructor and assemble packets for each student.
Code the packets and deliver to each subject.

Phase II (April 20-26) Administer Pretest.

- Step 1 Arrange for each teacher to administer the pretest for each class.
- Step 2 Collect the packets and deliver data to the computer processing center.

Phase III (May 4-10) Administration of Treatments.

- Step 1 Deliver pretest results to those instructors in pertinent treatment groups.
- Step 2 Videotape those instructors receiving videotape feedback for one hour during a class session mutually agreed upon by the instructor and the researcher.
- Step 3 Provide instructors who receive videotape feedback with videotape playback equipment for instructor self-viewing.

Phase IV (June 8-10) Administration of Posttest.

- Step 1 Assemble packets of response sheets and deliver to all thirty-six instructors.
- Step 2 Arrange for each instructor to administer the posttest for each class.
- Step 3 Collect the data and deliver data to the computer processing center for statistical analysis.

Summary

This chapter described research methodology used to meet objectives of the study.

The population of seventy-six full and part-time instructors teaching on-campus credit courses during the 1975 Spring quarter at New River Community College, Dublin, Virginia, was described. An equal-size stratified random sampling technique was described and five demographic characteristics of the population were identified.

The research design, a modification of the Solomon Four-Group design was described and a schema presented. Controls for threats to internal and external validity were enumerated. Three control groups and three treatment groups were described.

The instrument, Purdue Rating Scale for Instruction was discussed and items eliminated from the scale were listed.

The multivariate analysis of variance, linear contrasts, simultaneous confidence intervals, and multivariate T^2 test used to test the hypotheses were described.

Finally, the time schedule utilized to collect data and administer feedback was presented.

CHAPTER IV

PRESENTATION AND ANALYSES OF DATA

This chapter is a presentation and analyses of data collected as described in Chapter III.

Collection of Data

A total of 1,277 student ratings were collected from the thirty instructors receiving pretests and 1,595 were collected from the thirty-six instructors receiving the post-test.

Fifty-four of the 1,277 pretest ratings collected were affected by experimental mortality. These ratings were treated as described in Chapter III under the subheading, Experimental Mortality, and were not considered threats to internal validity.

Pretest and posttest student ratings collected for each treatment group by instructor is illustrated in the Appendix, Tables 13 and 14 respectively.

Computation of Gain Scores

Gain scores were computed in each treatment and control group by subtracting the mean of the pretest scores from the mean of the posttest scores for each item. Gain scores served as criteria for computing the statistical analyses. Gain scores are illustrated in the Appendix, Table 15.

Test of Hypothesis One

The test of hypothesis one, no significant difference in teaching behavior measured by student ratings shall occur between student ratings administered before and those administered after feedback is given the teacher about his teaching performance, utilized a multivariate T^2 test pairing the pretests and posttests of the nineteen instrument items. The computed T^2 value, presented by Table 4, was not sufficient to reject the null hypothesis at the .05 level. Therefore, no change in teacher behavior measured by student ratings occurred after feedback about teacher behavior was presented to teachers participating in the study.

Test of Hypothesis Two

The test of hypothesis two, no significant difference in teacher behavior measured by student ratings shall occur among teachers given different types of feedback, utilized a multivariate analysis of variance and the linear contrasts. The results are illustrated by Tables 5 and 6 respectively. The multivariate analysis of variance computed a Pillai's Trace of .0330 falling within the .05 alpha level for this study. However, the linear contrasts conducted by pairs of treatments revealed no statistically significant differences. The statistical treatment was not continued to compute the simultaneous confidence intervals since no difference existed between any pair of the treatment

TABLE 4

Multivariate T^2 Test for Hypothesis One

Statistical Computation	Computed Result	Parameters
Multivariate T^2 Test	*3.2462	df = 29 $\alpha = .05$

*Not significant at the .05 level

TABLE 5

Multivariate Analysis of Variance for Hypothesis Two

Statistical Computation	F Ratio Approximation	Parameters
Pillai's Trace	$F = 1.70626$ *Confidence Level = .0330	$S = 4$ $M = 7$ $N = 2,5$ $DF = 76,40$

*Denotes significance at the .05 level

TABLE 6

Linear Contrasts by Pairs of Treatments for
Hypothesis Two

	Significance Level
G1 - G2	0.1349
G1 - G3	0.7840
G1 - G4	0.3920
G1 - G5	0.4607
G2 - G3	0.2868
G2 - G4	0.0612
G2 - G5	0.1087
G3 - G4	0.4059
G3 - G5	0.3902
G4 - G5	0.2587

Legend: G1 = Combination Feedback
 G2 = Student Rating Feedback
 G3 = Control Group
 G4 = Videotape Feedback
 G5 = Second Control Group

groups. The null hypothesis was not rejected. Therefore, no difference in teacher behavior measured by student ratings occurred among teachers given different types of feedbacks.

Relationship of Demographic Characteristics to Teaching Behavior

To determine relationships between demographic characteristics and teaching behavior, a multivariate analysis of variance was conducted using strata of the characteristics as independent variables and gain scores of the instrument items as dependent variables. Approximate "F" values were computed for each item. Results were as follows:

Teaching Area. A relationship appeared to exist between teaching area strata, occupational/technical and college parallel, and three items of the instrument. Results of the analysis, illustrated in Table 7, revealed that three items, "sympathetic attitude toward students," "fairness in grading," and "stimulating intellectual curiosity," were related to teaching area.

Teaching Experience. A relationship appeared to exist between teaching experience strata, less than two years, two to four years, and more than four years, and three instrument items. Results of the analysis illustrated in Table 8 revealed that three items, "presentation of subject matter," "agreement between announced objectives of

TABLE 7

Multivariate Analysis of Variance Illustrating
Relationship Between Teaching Area and Instrument Items

Item Number	"F" Value Approximation	Confidence Level
1	1.30973	0.2919
2	2.52893	0.0401 *
3	3.84859	0.0060 *
4	0.96284	0.5025
5	1.56797	0.1920
6	0.50819	0.8516
7	0.49831	0.8586
8	0.29380	0.9679
9	0.59944	0.7838
10	2.51959	0.0407 *
11	0.93462	0.5178
12	1.43211	0.2396
13	0.79374	0.6267
14	0.66590	0.7305
15	0.51919	0.8438
16	0.89769	0.5452
17	0.68761	0.7129
18	0.30377	0.9644
19	0.61409	0.7719

*Denotes relationship at the .05 level

TABLE 8

Multivariate Analysis of Variance Illustrating
Relationship Between Teaching Experience and Instrument Items

Item Number	"F" Value Approximation	Confidence Level
1	0.41915	0.8918
2	1.71464	0.1751
3	1.05900	0.4391
4	1.10089	0.4147
5	4.13468	0.0089 *
6	2.31215	0.0769
7	0.64864	0.7278
8	0.99651	0.5226
9	0.30245	0.9532
10	0.30992	0.9500
11	0.72031	0.6729
12	0.42285	0.8895
13	0.44416	0.8759
14	3.76430	0.0131 *
15	1.51306	0.2326
16	2.34662	0.0734
17	1.72157	0.1734
18	2.48645	0.0609
19	2.63698	0.0501 *

*Denotes relationship at the .05 level

the course and what was actually taught," and "suitability of the amount and type of assigned outside work," were related to years of teaching experience.

Age. No relationship appeared to exist between age and instrument items. Results of the analysis are illustrated in Table 9.

Educational Level. No relationship appeared to exist between faculty education level and instrument items. Results of the analysis are illustrated by Table 10.

Occupational Experience Other Than Teaching. No relationship appeared to exist between years occupational experience and instrument items. Results of the analysis are illustrated by Table 11.

Summary

Test of Hypothesis One. Hypothesis one, no significant difference in teacher behavior measured by student ratings shall occur between student ratings administered before and those administered after feedback is given the teacher about his teaching performance, was tested at the .05 alpha level using a multivariate T^2 test. The hypothesis was not rejected.

Test of Hypothesis Two. Hypothesis two, no significant difference in teacher behavior measured by student ratings shall occur among teachers given different types of feedback, was tested at the .05 alpha level using a

TABLE 9

Multivariate Analysis of Variance Illustrating
Relationship Between Teacher Age and Instrument Items

Item Number	"F" Value Approximation	Confidence Level
1	1.37980	.2789
2	.83774	.5729
3	1.03975	.4432
4	.51193	.8131
5	1.30115	.3109
6	1.10342	.4071
7	.82026	.5852
8	.25287	.9631
9	1.63280	.1966
10	1.23313	.3413
11	.33025	.9283
12	.79735	.6016
13	.26535	.9582
14	.67595	.6914
15	1.17474	.3697
16	.63213	.7243
17	.19305	.9818
18	.10170	.9966
19	.38584	.8971

*Denotes relationship at the .05 level

TABLE 10

Multivariate Analysis of Variance Illustrating
Relationship Between Education Level and Instrument Items

Item Number	"F" Value Approximation	Confidence Level
1	.40136	.8066
2	1.39018	.2745
3	2.81196	.0542
4	1.22060	.3349
5	.84482	.5156
6	.72505	.5877
7	.22992	.9169
8	1.00713	.4295
9	2.30636	.0950
10	.86574	.5038
11	.34317	.8459
12	1.09754	.3868
13	.52374	.7220
14	2.39466	.0860
15	1.26526	.3178
16	2.34183	.0913
17	2.01834	.1321
18	1.78599	.1728
19	1.92148	.1477

*Denotes relationship at the .05 level

TABLE 11

Multivariate Analysis of Variance
 Illustrating Relationship Between Occupational Experience
 Other Than Teaching and Instrument Items

Item Number	"F" Value Approximation	Confidence Level
1	0.49955	0.8010
2	0.84931	0.5508
3	1.95748	0.1289
4	0.64840	0.6926
5	1.21672	0.3452
6	0.48445	0.8116
7	0.26924	0.9431
8	1.17664	0.3640
9	0.93836	0.5058
10	1.24068	0.3344
11	0.53358	0.7765
12	0.46610	0.8245
13	1.48992	0.2400
14	0.83329	0.5615
15	0.73009	0.6333
16	1.37193	0.2809
17	0.59521	0.7316
18	0.93648	0.5046
19	2.24087	0.0889

*Denotes relationship at the .05 level

multivariate analysis of variance and linear contrasts. The hypothesis was not rejected.

Relationship of Demographic Characteristics to Teaching Behavior. A one-way multivariate analysis of variance was used to compute approximate "F" tests to determine relationships of demographic characteristics to instrument items. Three items, "sympathetic attitude toward students," "fairness in grading," and "stimulating intellectual curiosity" appeared to be related to teaching area.

Three items, "presentation of subject matter" agreement between announced objectives of the course and what was actually taught," and, "suitability of the amount and type of assigned outside work," appeared to be related to years teaching experience.

No items appeared to be related to age of the instructor, educational level of the instructor, and occupational experience other than teaching.

CHAPTER V

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

Summary

Efforts to determine what constitutes effective teaching have not been productive. No widely accepted definition of effective teaching nor validated system for evaluating effective teaching has been established. An evaluation scheme providing administrators and supervisors with objective and quantitative data about teacher performance appears needed. The scheme should provide instructors with feedback about teaching behavior and enhance opportunities to modify undesirable behavior before data are presented to supervisors.

Statement of the Problem. The problem of this study was concerned with utilization of feedback which enabled teachers to modify teaching behavior after receiving the feedback. Objectives of the study were:

1. To provide behavioral feedback to instructors and statistically measure resultant behavioral changes.
2. To review current literature and report types of feedback used.
3. To select three types of feedback and statistically measure which types caused behavioral changes.

4. To choose a validated student rating instrument identifying and measuring aspects of teaching behavior.

5. To select five demographic characteristics of the population and determine their relationship to teaching behavior.

Statement of Hypotheses. 1. No significant difference in teaching behavior measured by student ratings shall occur between student ratings administered before and those administered after feedback is given the teacher about his teaching performance.

2. No significant difference in teaching behavior measured by student ratings shall occur among teachers given different types of feedback.

Research Methodology. A stratified random sample of thirty-six instructors were drawn from a population of seventy-six full-time and part-time instructors employed at New River Community College, Dublin, Virginia, during the Spring quarter of the academic year, 1974-75. Six instructors were assigned to each of two control groups, three treatment groups, and a posttest-only group, using an equal-size stratified random sampling technique. The research design was a modified Solomon Four-Group.

A pretest and posttest, consisting of nineteen items of the Purdue Rating Scale for Instruction, was administered to all groups except the posttest-only group.

Gain scores (posttest scores minus pretest scores) were computed for each group.

Test of the Hypotheses. Hypothesis one, no significant difference in teaching behavior measured by student ratings shall occur between student ratings administered before and those administered after feedback is given the teacher about his teaching performance, was tested at the .05 alpha level using a multivariate T^2 test. The hypothesis was not rejected.

Hypothesis two, no significant difference in teaching behavior measured by student ratings shall occur among teachers given different types of feedback, was tested at the .05 alpha level using a multivariate analysis of variance and linear contrasts. The hypothesis was not rejected.

Relationship of Demographic Characteristics To Teaching Behavior. Approximate "F" tests were conducted to determine relationships between demographic characteristics and each of the nineteen instrument items.

1. Teaching area. Three items, "sympathetic attitude toward students," "fairness in grading," and "stimulating intellectual curiosity" appeared to be related to teaching area.

2. Teaching experience. Three items, "presentation of subject matter," "agreement between announced objectives

of the course and what was actually taught," and "suitability of the amount and type of assigned outside work," appeared to be related to years of teaching experience.

3. Age. No items appeared to be related to age of instructors.

4. Educational level. No items appeared to be related to educational level.

5. Years occupational experience other than teaching. No items appeared to be related to years of occupational experience other than teaching.

Results

The following were results of the study:

1. No significant change occurred in teaching behavior as measured by student ratings after feedback was administered.

2. No specific type of feedback was associated with behavioral changes.

3. No significant change occurred in teaching behavior during the ten-week quarter.

4. There appeared to be no relationship between age of instructors, educational level of instructors and occupational experience other than teaching and instrument items.

5. There appeared to be a relationship between teaching area and three instrument items, "sympathetic

attitude toward students," "fairness in grading," and "stimulating intellectual curiosity."

6. There appeared to be a relationship between teaching experience and three instrument items, "presentation of subject matter," "agreement between announced objectives of the course and what was actually taught" and "suitability of the amount and type of assigned outside work."

Discussion of Results

Reflection upon results of the study provided a basis for the following conclusions:

1. No change in teaching behavior occurred after feedback was administered. This appeared to be the result of several circumstances. Pretest scores were higher than anticipated, averaging between 4.0 and 4.5. The greatest gain score was only .5183, illustrating a minor change. High pretest scores left little room for improvement. The high pretest scores at New River Community College might be attributed to a uniqueness of that institution's population.

Veteran teachers appeared less inclined to make behavioral changes. Veteran teachers were equally dispersed in the treatment groups thereby neutralizing more pronounced changes by neophyte teachers.

Administration of video feedback did not include a description of specific behavior to which teachers should have been alerted. The possibility existed that teachers

not familiar with viewing themselves failed to observe undesirable teaching behavior.

The study was conducted during Spring quarter after students had become familiar with teachers' behaviors. This familiarity might have caused students to overlook undesirable teaching behavior.

The study utilized a research design controlling unwanted variables. Some of these variables are normally present during the teaching process. Omission of these variables might have eliminated aspects of the teaching process which result in behavioral changes.

No correlation was made to determine if students and teachers agreed what constitutes positive and negative teaching behavior. This researcher believes there may not have been a commonality of students' perceptions of poor teaching behavior and teachers' perceptions of poor teaching behavior.

2. No specific type of feedback was associated with behavioral changes. This appeared contrary to significant findings reported in the literature review (14), (30), (35), (54), (56), (66). This researcher expected the combination of videotaping and student ratings to result in behavioral changes. That changes did not occur might have been failure of teachers to utilize the two sources of feedback together. Separate utilization of feedback may have negated effects of combination treatment.

The video feedback and student feedback might have failed to reveal the same negative teaching behaviors. This researcher believes the combination of two feedback sources did not reinforce each other but acted as separate feedbacks.

3. No demographic characteristics were significantly related to teaching behavior. This result was anticipated. This anticipation led the researcher to stratify the sample.

4. No change in teaching behavior occurred during the ten-week quarter. This appeared to be related to the limited time for teachers to receive feedback and change behavior. Four weeks was the total time allowed for changes to occur. Only one opportunity to receive feedback was provided each teacher in the study. Feedback opportunities provided during several quarters might result in behavioral changes.

Recommendations

Based on results it was recommended that:

1. Administrators of community colleges and similar institutions investigate possibilities that demographic characteristics of instructors are related to results of student ratings.

2. Administrators and instructors of community colleges and similar institutions identify ways to quantify teaching behavior other than student ratings.

3. Instructors at community colleges and similar institutions determine if initial student ratings are skewed positively making subsequent positive changes in teaching behavior difficult to measure.

4. A similar study be conducted using more than one academic quarter to conduct the research.

5. A similar study be conducted during the Fall academic quarter using ratings from freshman students.

6. Research be conducted which statistically tests effects of teacher demographic characteristics on teaching behavior.

7. A similar study be conducted using trained researchers to administer video feedback.

8. Student rating instruments be developed which measure teaching behavior which can be recorded by a video camera.

9. Research be conducted which compares effects of feedback and demographic characteristics of teachers on student ratings.

10. A similar study be conducted giving teachers more than one opportunity to receive feedback.

11. Research be conducted to correlate students' perceptions of positive teaching behavior with teachers' perceptions of positive teaching behavior.

12. A similar study be conducted using an alternative student rating instrument measuring teaching behavior.

13. This study be replicated in one or more community colleges to assess possibilities that the population of New River Community College might have been unique with respect to student evaluation of faculty.

REFERENCES

REFERENCES

1. Abramson, P. (ed.) "Sauce for the Goose Is Sauce for the Gander." College Management, July, 1966. pp. 27-31.
2. Aleamoni, Lawrence M. "The Usefulness of Student Evaluations in Improving College Teaching." Research Report no. 356. Illinois University Urbana Office of Instructional Resources, 1974.
3. Allen, Dwight W., and Frederick J. McDonald. "Training Effects of Feedback and Modeling Procedures on Teaching Performance." Unpublished Doctor's dissertation, Stanford University, California, 1967.
4. Amidon, Edmund J., and Barak Rosenshine. "Interaction Analysis and Microteaching in an Urban Teacher Education Program. A Model for Skill Development in Teaching." Paper presented at the American Educational Research Association convention, February, 1968, Chicago, Illinois.
5. Armstrong, Harold R. "A Teacher's Guide to Teaching Performance Evaluation." School Management Institute, Worthington, Ohio, 1972.
6. Baird, J. Hugh, and others. "Micro-teaching at Brigham Young University." Available from ERIC Document Service (ED 011260), February, 1967.
7. Barney, Jacqueline D. "Evaluation of Teaching Through Utilization of Videotapes: Preservice and Inservice." Dover, Delaware: Delaware Department of Public Instruction, Division of Occupational Research, 1971.
8. Bayley, D.H. "Making College Teaching a Profession." Improving College and University Teaching, 15:115-119, June, 1967.
9. Belt, W. Dwayne. "Micro-Teaching - Observed and Critiqued By a Group of Trainees." Paper read at the American Educational Research Association annual meeting, February, 1967, New York.
10. Bhaerman, Robert D. "Merit Pay? No," National Elementary Principal, 52:63-68, February, 1973.

11. Bierschenk, Bernard. "Self-Confrontation Via Closed-Circuit Television in Teacher Training: Results, Implications and Recommendations." Available from ERIC Document Service (ED070293), September, 1972.
12. Birnbaum, R. "Background and Evaluation of Faculty in New York." The Junior College Journal, 38:34-37, 1966.
13. Bolin, John G. and John W. Muir. "Merit Rating for Salary Increases and Promotions." Institute of Higher Education, University of Georgia, Athens, Georgia, 1966.
14. Borg, Walter R. and others. "The Effects of Videotape Feedback and Microteaching in a Teacher Training Model." Berkeley, California, Far West Laboratory for Higher Educational Research and Development, 1968.
15. Borgatta, E.F. "Student Rating of Faculty," AAUP Bulletin, 56:6-7, 1970.
16. Brickman, W.W. "Student Evaluation of Professors." School and Society, 94:143-144, 1966.
17. Brown, D.W. "Teach or Perish." Improving College and University Teaching, 15:108-110, 1967.
18. Callahan, Leroy G. "Clinical Evaluation and the Classroom Teacher." Paper read at the American Educational Research Association meeting.
19. Cameron, Walter A., and Calvin J. Cotrell. "Remote Feedback Techniques for Inservice Education. Assessment of Micro-Teaching and Video Recording in Vocational and Technical Teacher Education: Phase X-Interim Report. Research and Development Series No. 40." Columbus, Ohio, Ohio State University, The Center for Vocational and Technical Education, July, 1970.
20. Centra, J.A. "The Effectiveness of Student Feedback in Modifying College Instruction." The Utility of Student Ratings for Instructional Improvement, Princeton: Educational Testing Service, 1972.

21. Clark, S.J., and P.L. Miller. "Student Evaluation of Teachers-Is It Useful, Consistent, and Honest?" Speech given at the annual meeting of the American Society for Engineering Education, June 22, 1971.
22. Clayton, Thomas E., and others. "Preparing Beginning Teachers for Working with the Educationally Disadvantaged: A Continuing Study of the Influence on Student Teachers of a Training Program for Cooperating Teachers." Available from ERIC Document Service (ED 036460), July, 1969.
23. Codwell, John E. "A Demonstration of the Effect of an Adaptation of Microteaching on the Instructional Behavior of Rural School Teachers." Available from ERIC Document Service (ED 034620), October, 1969.
24. Cohen, Arthur M. "Evaluation of Faculty." Community College Review, 2:12-21, Summer, 1974.
25. Costin, F., and others. "Student Ratings of College Teaching: Reliability, Validity, and Usefulness." Review of Educational Research, 41:511-535, 1971.
26. Dugas, Donald G. "Micro-Teaching - A Promising Medium for Teacher Training." The Modern Language Journal, 51:161-166, March, 1967.
27. Dunay, Littian R., and Helen F. Robison. "Student Teachers' Self Analysis of Teaching Behavior." Paper read at the American Educational Research Association meeting, February, 1973, New Orleans, Louisiana.
28. Friebel, Allen C., and Warren W. Kallenbach. "Effects of Videotape Feedback and Microteaching as Developed in the Field Test of Minicourse I with Student Teachers." Paper read at the California Educational Research Association meeting, March, 1969, Los Angeles, California.
30. Gage, N.L., and others. "Equilibrium Theory and Behavior Change: An Experiment in Feedback from Pupils to Teachers." Available from ERIC document Service (ED 049176), August, 1960.
31. Gall, Meredith D., and others. "The Relationship Between Personality and Teaching Behavior Before and After Inservice Microteaching Training." Available from ERIC Document Service (ED 031448), March, 1969.

32. Greenberg, Jerrold S. "How Videotaping Improves Teaching Behavior." Jomper, 23:36-37, March, 1973.
33. Harrington, Dred W. "Assessment of Micro-Teaching and Video Recording in Vocational and Technical Education: Feedback Techniques for In-Service Technical Teacher Education." Center for Vocational and Technical Education, Ohio State University, 1971.
34. Hasse, Richard F., and C. Dean Miller. "Student Evaluation of Teachers' Competence and Effectiveness." Paper read at the American Personnel and Guidance Association Convention, March, 1967, Dallas, Texas.
35. Hoehn, Lilburn P. (ed.). Teaching Behavior Improvement Program. Available from Michigan-Ohio Regional Educational Laboratory, Inc., 3750 Woodward Avenue, Detroit, Michigan, July, 1969.
36. Hovenier, Peter J. "Changing the Behavior of Social Studies Department Heads Through the Use of Feedback." Available from ERIC Document Service (ED 010436), August, 1966.
37. Hrivnak, Joseph T. "The Use of Interaction and Feedback in an Inservice Education Model." Unpublished Doctor's dissertation, University of Pittsburgh, 1970.
38. Johnson, Mel. "Model Program for Teacher In-Service Training Emphasizing the Affective Dimension." Available from Elk Grove Training and Development Center, Arlington Heights, Illinois, June, 1969.
39. Kallenbach, Warren W. "Microteaching as a Teaching Methodology." Paper read at the Conference on Instructional Methods and Teacher Behavior, November 21, 1966, Berkeley, California.
40. Karasar, Niyazi. "Impact of Video Feedback on Teachers' Eye-Contact Mannerisms in Microteaching." Unpublished Doctor's dissertation, Ohio State University, 1970.
41. Langer, Philip. "Minicourse-Theory and Strategy." Paper read at the annual meeting of the American Educational Research Association, February 7, 1969, Los Angeles, California.

42. Larson, Eleanore A., and Thomas A. Ringness. "Effect of Clinical Evaluation on Supervision of Student Teachers." Available from ERIC Document Service (ED 003467), 1965.
43. Lauroesch, William P., and others. "The Use of Student Feedback in Teacher Training." Unpublished Doctor's dissertation, Chicago University, June, 1969.
44. McDonald, Frederick J. "A Behavior Modification View of Video Playback: Microteaching." Paper read at the American Educational Research Association meeting, 1973, New Orleans, Louisiana.
45. Meers, Gary D. "Effects of Various Feedback Media on Psychomotor Achievement." Unpublished Doctor's dissertation, Missouri University, Columbia, Missouri, May, 1972.
46. Miller, M.T. "Instructor Attitudes Toward, and Their Use of, Student Ratings of Teachers." Journal of Educational Psychology, 62:235-239, 1971.
47. Moe, Alden J., and Sister Mary Dorothy Feehan. "The Use of Videotape Recorders in the Training of Reading Teachers." Paper read at the International Reading Association conference, April 24, 1968, Boston, Massachusetts.
48. Morse, Kevin R., and others. "Effects of Different Types of Supervisory Feedback on Teacher Candidates' Development of Refocusing Behaviors." Available from ERIC Document Service (ED 046906), September, 1970.
49. Naeslund, Jon. "Experiments with Closed Circuit Television." Available from ERIC Document Service (ED 073631), November, 1972.
50. Newell, D. "Evaluation of Teachers." Paper read at the College of Dentistry Conference on Evaluation of Teaching and Teachers, 1967, University of Kentucky, Lexington, Kentucky.
51. Northfield Project. "Carleton College Use of Video Tapes in Training Secondary School Teachers." Available from ERIC Document Service (ED 011245), November, 1966.

52. Novak, John H. "A Study of the Effects of the Use of a Pupil Response Instrument on the Behaviors of Biological Science Teachers." Available from ERIC Document Service (ED 069155), November, 1972.
53. Oliver, Wilmot F. "The Relative Effectiveness of Informational Feedback About Supervisory and Student Reactions with Beginning and Experienced Vocational Teachers." Available from ERIC Document Service (ED 019477), October 1967.
54. Pachla, Mark. "Feedback Expectancy and Shifts in Student Ratings of College Faculty." Journal of Applied Psychology, 58:254-258, October, 1973.
55. Pambookian, Hogop S. "Initial Level of Student Evaluation of Instruction as a Source of Influence on Instructor Change After Feedback." Journal of Educational Psychology, 66:52-56, December, 1974.
56. Perlberg, Arye, and David C. O'Bryant. "The Use of Videotape Recording and Microteaching Techniques to Improve Instruction on the Higher Education Level." Available from ERIC Document Service (ED 023314), August, 1968.
57. Redfern, George B. "Will Teacher Militancy Make Evaluation of Teacher Performance Obsolete?" Paper read at the annual meeting of the American Association of School Administrators, February 17, 1969, Atlantic City, New Jersey.
58. Ronan, W.W. "Development of an Instrument to Evaluate College Classroom Teaching Effectiveness." Unpublished Doctor's dissertation, Georgia Institute of Technology, 1971.
59. Roush, Robert E. "A Study of Change in Selected Teacher Education Interns' Behavior Using Videotape Recordings." Available from ERIC Document Service (ED 031095), January 1969.
60. Salomon, Gavriel, and Frederick J. McDonald. "Pre- and Posttest Reactions of Self-Viewing Ones Performance on Videotape." Paper read at the annual convention of the American Psychological Association, August 30, 1968, San Francisco, California.

61. Stukat, Karl-Gustaf. "Microteaching." Available from ERIC Document Service (ED 073630), November 1972.
62. Urbach, Floyd D., and others. "Instructional Staff Development. Component 2: Inquiry Influence." Available from Mid-Continent Regional Educational Laboratory, Inc., Kansas City, Missouri, July, 1972.
63. Wagner, A.C. "Changing Teaching Behavior: A Comparison of Microteaching and Cognitive Discrimination Training." Journal of Educational Psychology, 64:299-305, June, 1973.
64. Waimon, Morton D. and Gary C. Ramseyer. "Effectives of Video Feedback on the Ability to Evaluate Teaching." The Journal of Teacher Education, 21:92-95, Spring, 1970.
65. Ward, Phillip M. "The Use of the Portable Videotape Recorder in Helping Teachers Self-Evaluate Their Teaching Behavior." Available from ERIC Document Service (ED 038365), January, 1970.
66. Webb, Clark, and others. "Description of a Large Scale Micro-Teaching Program." Paper read at the Department of Audiovisual Instruction National Conference, March, 1968, Houston, Texas.
67. Wise, Charles N. "Student Ratings of Teachers: A Perspective for Speech Communication." Western Speech, 23:196-203, Summer, 1973.
68. Young, David B. "Effective Supervisory Conferences: Strategies for Modifying Teacher Behavior." Available from ERIC Document Service (ED 041840), 1969.
69. Young, Dorothy A., and David B. Young. "The Effectiveness of Individually Prescribed Micro-Teaching Training Modules on an Intern's Subsequent Classroom Performance." Paper read at the American Educational Research Association annual meeting, February, 1969, Los Angeles, California.

APPENDIX

TABLE 12

THE PURDUE RATING SCALE FOR INSTRUCTION

By H. H. Klemm and D. N. Elliott

Note to Instructors: To keep conditions as nearly uniform as possible, it is imperative that no instructions be given to the students. The rating scale should be passed out without comment at the beginning of the period.

Note to Students: Following is a list of qualities that, taken together, tend to make any instructor the sort of instructor that he is. Of course, no one is ideal in all of these qualities, but some approach this ideal to a much greater extent than do others. To provide information which may lead to the improvement of instruction, you are asked to rate your instructor on the indicated qualities by darkening one of the spaces on the line at the point which most nearly describes him with reference to the quality you are considering. For example, under Interest in Subject if you think your instructor is not as enthusiastic about his subject as he should be, but is usually more than mildly interested, darken the space marked 6D on the answer card. Fill the chosen space solidly with a special electrographic or soft lead pencil; leave no stray marks.

6. Interest in Subject _____ 6A 6B 6C 6D 6E 6F 6G 6H 6I 6J
Always appears full of his subject. Seems mildly interested. Subject seems irksome to him.

DO NOT WRITE ON THIS PAGE—MARK ALL RESPONSES ON THE ANSWER CARDS.

This rating is to be entirely impersonal. Do not sign your name or make any other mark on the paper which could serve to identify the rater.

1. Interest in Subject _____	1A	1B	1C	1D	1E	1F	1G	1H	1I	1J
	Always appears full of his subject.			Seems mildly interested.			Subject seems irksome to him.			
2. Sympathetic Attitude toward Students _____	2A	2B	2C	2D	2E	2F	2G	2H	2I	2J
	Always courteous and considerate.			Tries to be considerate but finds it difficult at times.			Entirely unsympathetic and inconsiderate.			
3. Fairness in Grading _____	3A	3B	3C	3D	3E	3F	3G	3H	3I	3J
	Absolutely fair and impartial to all.			Shows occasional favoritism.			Constantly shows partiality.			
4. Liberal and Progressive Attitude _____	4A	4B	4C	4D	4E	4F	4G	4H	4I	4J
	Welcomes differences in viewpoint.			Biased on some things but usually tolerant.			Entirely intolerant, allows no contradiction.			
5. Presentation of Subject Matter _____	5A	5B	5C	5D	5E	5F	5G	5H	5I	5J
	Clear, definite and forceful.			Sometimes mechanical and monotonous.			Indefinite, involved, and monotonous.			
6. Sense of Proportion and Humor _____	6A	6B	6C	6D	6E	6F	6G	6H	6I	6J
	Always keeps proper balance; not over-critical or over-sensitive.			Fairly well balanced.			Over-serious; no sense of relative values.			
7. Self-reliance and Confidence _____	7A	7B	7C	7D	7E	7F	7G	7H	7I	7J
	Always sure of himself; meets difficulties with poise.			Fairly self-confident; occasionally disconcerted.			Hesitant, timid, uncertain.			
8. Personal Peculiarities _____	8A	8B	8C	8D	8E	8F	8G	8H	8I	8J
	Wholly free from annoying mannerisms.			Moderately free from objectionable peculiarities.			Constantly exhibits irritating mannerisms.			
9. Personal Appearance _____	9A	9B	9C	9D	9E	9F	9G	9H	9I	9J
	Always well groomed; clothes neat and clean.			Usually somewhat untidy; gives little attention to appearance.			Slovenly; clothes untidy and ill-kept.			
10. Stimulating Intellectual Curiosity _____	10A	10B	10C	10D	10E	10F	10G	10H	10I	10J
	Inspires students to independent effort; creates desire for investigation.			Occasionally inspiring; creates mild interest.			Destroys interest in subject; makes work repulsive.			

Note to Students: Following is a list of factors which are important to many courses but over which the instructor often has little control. You are asked to rate the course on each of the factors by selecting one of the spaces at the right of each statement and darkening the corresponding space on the answer card. Use a special electrographic or soft lead pencil. MAKE NO MARKS ON THIS PAGE. RECORD ALL RESPONSES ON YOUR ANSWER CARD.

If the course is extremely poor with respect to the factor darkened space E:

A B C D E

If the course is below average with respect to the factor darkened space D:

A B C D E

If the course is average with respect to the factor darkened space C:

A B C D E

If the course is above average with respect to the factor darkened space B:

A B C D E

If the course is excellent with respect to the factor darkened space A:

A B C D E

For example: If you feel that the course is not contributing very much to the attainment of your ultimate goal, but on the other hand, is not a complete waste of time you would probably respond to item number 20 by darkening space 20D on your answer card.

20. How the course is fulfilling your needs (consider your ultimate as well as your immediate goals) 20A 20B 20C 20D 20E

ANSWER CARD

18A	19A	20A	21A	22A
18B	19B	20B	21B	22B
18C	19C	20C	21C	22C
18D	19D	20D	21D	22D
18E	19E	20E	21E	22E

11. Suitability of the method or methods by which subject matter of the course is presented (recitation, lecture, laboratory, etc.)

11A 11B 11C 11D 11E

12. Suitability of the size of the class (consider the subject matter and type of class—lecture, lab, etc.)

12A 12B 12C 12D 12E

13. The degree to which the objectives of the course were clarified and discussed

13A 13B 13C 13D 13E

14. The agreement between the announced objectives of the course and what was actually taught

14A 14B 14C 14D 14E

15. Suitability of the reference materials available for the course

15A 15B 15C 15D 15E

16. Suitability of the laboratory facilities available for the course

16A 16B 16C 16D 16E

17. Suitability of the assigned textbook

17A 17B 17C 17D 17E

18. The use made of tests as aids to learning

18A 18B 18C 18D 18E

19. Amount of freedom allowed students in the selection of the materials to be studied (considering the subject matter)

19A 19B 19C 19D 19E

20. How the course is fulfilling your needs (consider your ultimate as well as your immediate goals)

20A 20B 20C 20D 20E

21. Range of ability in the class (are there too many extremely dull or extremely bright students?)

21A 21B 21C 21D 21E

22. Suitability of the amount and type of assigned outside work

22A 22B 22C 22D 22E

23. The weight given to tests in determining the final grade for the course

23A 23B 23C 23D 23E

24. Coordination of the tests with the major objectives of the course

24A 24B 24C 24D 24E

25. Frequency of tests

25A 25B 25C 25D 25E

26. The overall rating of the instructor

26A 26B 26C 26D 26E

TABLE 12 (continued)

TABLE 13

Number of Student Ratings Collected for Pretest

INSTRUCTORS	GROUPS				
	GROUP #1 Combination Feedback	GROUP #2 Student Rating Feedback	GROUP #3 Control	GROUP #4 Video Feedback	GROUP #5 Second Control
1	41	54	91	107	8
2	51	19	42	32	47
3	15	54	26	4	18
4	24	21	11	56	42
5	33	20	86	117	26
6	47	25	74	48	19
Totals	211	193	330	364	160

Total Pre-Evaluation Ratings Collected - 1277

TABLE 14

Number of Student Ratings Collected for Posttest

INSTRUCTORS	GROUPS					
	GROUP #1 Combination Feedback	GROUP #2 Student Rating Feedback	GROUP #3 Control	GROUP #4 Video Feedback	GROUP #5 Second Control	GROUP #6 Posttest Only
1	41	54	90	104	8	15
2	51	19	42	32	47	13
3	15	52	26	4	18	20
4	20	21	6	56	42	22
5	33	20	86	106	26	171
6	41	23	74	48	19	131
Totals	201	188	324	350	160	372

Total Post-Evaluation Ratings Collected - 1595

TABLE 15

Gain Scores by Item for Each Group

ITEM NUMBER	GROUPS				
	GROUP #1 Teaching Area	GROUP #2 Teaching Experience	GROUP #3 Age	GROUP #4 Education Level	GROUP #5 Occupational Experience
1	-.0133	-.0767	.0083	.1283	-.1950
2	.0983	.0950	-.0017	.1217	-.2683
3	.0483	.0833	-.2083	.0767	-.4850
4	.0567	.1167	.0450	-.0117	-.1417
5	.3933	.0467	.1600	-.0150	-.0167
6	.1233	.1417	.0183	.0183	-.0800
7	.1633	.0950	.0667	.0433	-.0317
8	.0967	.2633	.1167	.0567	.0617
9	.0367	-.0617	.0900	.0800	-.1400
10	.0883	.0150	.1483	.0367	-.2167
11	.2183	.4617	.1233	-.2183	.0600
12	.2150	.3550	.1900	.0483	.1950
13	-.0667	.3000	.0333	.1550	-.0433
14	.1300	.2533	.0400	.1017	.0333
15	.1117	.1817	.0500	.2383	.0983
16	.0583	.3733	.0817	.1367	-.0350
17	.2250	.5083	.0533	.2600	.4983
18	.2300	.3533	.1117	.2967	.5183
19	.2217	.4200	.1383	.1483	.2367

**The two page vita has been
removed from the scanned
document. Page 1 of 2**

**The two page vita has been
removed from the scanned
document. Page 2 of 2**

A PILOT STUDY TESTING EFFECTS OF
FEEDBACK ON TEACHING BEHAVIOR AS
MEASURED BY STUDENT RATINGS AT
NEW RIVER COMMUNITY COLLEGE,
DUBLIN, VIRGINIA

by

Norman H. Scott

(ABSTRACT)

Research Problem. The focus of this study was utilization of feedback which enabled teachers to modify teaching behavior after receiving the feedback.

Research Methodology. The population was seventy-six full time and part time instructors employed at New River Community College, Dublin, Virginia, during the Spring quarter of the academic year, 1974-75. Six instructors were assigned to each of two control groups, three treatment groups, and a posttest-only group, using an equal-size stratified random sampling technique. The research design was a modified Solomon Four-Group.

Pretests and posttests (Purdue Rating Scale for Instruction) were administered to all groups except the posttest-only group.

Test of the Hypotheses. Data collected from the pretests and posttests were used to test two hypotheses. Hypothesis one, no significant difference in teaching behavior measured by student ratings shall occur between

student ratings administered before and those administered after feedback is given the teacher about his teaching performance, was tested at the .05 alpha level using a multivariate T^2 test. The hypothesis was not rejected.

Hypothesis two, no significant difference in teaching behavior measured by student ratings shall occur among teachers given different types of feedback, was tested at the .05 alpha level using a multivariate analysis of variance and linear contrasts. The hypothesis was not rejected.

Results.

1. No significant change occurred in teaching behavior as measured by student ratings after feedback was administered.

2. No specific type of feedback was associated with behavioral changes.

3. No significant change occurred in teaching behavior during the ten-week quarter.

4. There appeared to be no relationship between age of instructors, educational level of instructors and occupational experience other than teaching and any one of the instrument items of the Purdue Rating Scale.

5. There appeared to be a relationship between teaching area and three instrument items.

6. There appeared to be a relationship between teaching experience and three instrument items.

Conclusions.

1. Pretest scores were higher than anticipated apparently making subsequent posttest scores not significantly different.
2. Student familiarity achieved after a two-quarter exposure to teachers might have resulted in students overlooking undesirable teaching behavior.
3. Failure to correlate students' and teachers' perceptions of teaching behavior may have eliminated commonalities needed for teachers to make self-assessments using student ratings.
4. Lack of association of specific feedback with behavioral changes appeared to result from failure of teachers to utilize the combination feedback appropriately.
5. One ten-week quarter appeared insufficient time for teachers to change behavior needing change.

Recommendations. Based on conclusions it was recommended that research be conducted investigating:

1. Possibilities that demographic characteristics of instructors are related to results of student ratings.
2. The likelihood that student ratings are skewed positively making subsequent positive changes in teaching behavior difficult to measure.
3. The use of more than one academic quarter to conduct the research.
4. Effects of teacher demographic characteristics on teaching behavior.

5. Comparison of effects of feedback and demographic characteristics of teachers on student ratings.

6. The correlation between teachers' and students' perceptions of positive teaching behavior.