HOLLERAN, Brian Patrick, 1942-
EFFECTS OF PUPIL FEEDBACK UPON THE
COMMUNICATION BEHAVIORS OF STUDENT
TEACHERS.

The Ohio State University, Ph.D., 1969
Education, teacher training

University Microfilms, Inc., Ann Arbor, Michigan

Copyright
Brian Patrick Holleran
1970
EFFECTS OF PUPIL FEEDBACK UPON THE COMMUNICATION BEHAVIORS OF STUDENT TEACHERS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in the Graduate School of The Ohio State University

By

Brian Patrick Holleran, B.S., M.A.

* * * *

The Ohio State University
1969

Approved by

[Signature]
Adviser
College of Education
ACKNOWLEDGMENTS

The writer wishes to publicly express his appreciation to the many people who contributed to the development and completion of this dissertation.

To his major adviser, Professor George L. Lewis, for his inspiration, patience, and able counseling.

To his reading committee, Professor James Golden, Professor L.O. Andrews, and Professor Herbert Coon, for their constructive criticism and encouragement.

To Theresa Fecheck, Edward Long, and Joseph Abuscato, who kindly cooperated in collection of data from the student teachers they supervised.

To all the student teachers who volunteered to participate in the study and return the results of the pupil evaluations.

To Linda Vincent, who painstakingly transferred data to IBM cards.

And finally, to my sister, Terrie, who patiently struggled through the translation of my hieroglyphics and competently typed the manuscript.
VITA

August 30, 1942 Born - Youngstown, Ohio
1960-1961 ..... Attended John Carroll University, Cleveland, Ohio
1961-1964 ..... Attended Kent State University, Kent, Ohio
1964 .......... B.S. in Education, Kent State University, Kent, Ohio
1964-1966 ..... Teaching Assistant, The Ohio State University, Columbus, Ohio
1966 .......... M.A., The Ohio State University, Columbus, Ohio
1966-1967 ..... Teacher in public high school, Westerville, Ohio
1967-1969 ..... Teaching Associate, The Ohio State University, Columbus, Ohio

FIELDS OF STUDY

Major Field: Education

B.S. Secondary Education
M.A. General Communications
Ph.D.

  Studies in Speech Education. Professor George L. Lewis

  Studies in Rhetoric and Public Address. Professor James L. Golden

  Studies in Teacher Education. Professor L.O. Andrews
## CONTENTS

<table>
<thead>
<tr>
<th>ACKNOWLEDGMENTS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ii</td>
</tr>
<tr>
<td>VITA</td>
<td>iii</td>
</tr>
<tr>
<td>TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>ILLUSTRATIONS</td>
<td>viii</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
</tr>
<tr>
<td>BACKGROUND OF THE PROBLEM</td>
</tr>
<tr>
<td>Statement of the Problem</td>
</tr>
<tr>
<td>Feedback and Learning Theory</td>
</tr>
<tr>
<td>Feedback and Equilibrium Theory</td>
</tr>
<tr>
<td>The Instrument of Feedback</td>
</tr>
<tr>
<td>Questions and Hypotheses to be Studied</td>
</tr>
<tr>
<td>Design of the Study</td>
</tr>
<tr>
<td>Limitations of the Study</td>
</tr>
<tr>
<td>Definitions of Terms</td>
</tr>
<tr>
<td>Significance of the Study</td>
</tr>
<tr>
<td>The Plan of the Dissertation</td>
</tr>
</tbody>
</table>

| II       |
| REVIEW OF THE THEORY AND LITERATURE | 24 |
| Theories of Communication | 24 |
| Theories of Learning | 30 |
| Theories of Equilibrium | 32 |
| Integration of Theories | 34 |
| Values of Pupil Evaluation of Teachers | 38 |
| Studies in Teacher Evaluation | 44 |
| Studies in Pupil Evaluation of Student Teachers | 49 |
| Summary of Studies in Teacher Evaluation | 53 |

<p>| III      |
| DESIGN OF THE STUDY | 55 |
| Population Sample | 55 |
| Development of the Feedback Form | 58 |
| Processes Used in Interpreting the Data | 63 |</p>
<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PRESENTATION AND STATISTICAL ANALYSIS OF DATA</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV</td>
<td>Pupils' Conception of Their Learning</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td>Classification of Learnings</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>Statistical Data on the Null Hypotheses</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>Hypothesis 1</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>Hypothesis 2</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>Hypothesis 3</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>Hypothesis 4</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>Hypothesis 5</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>Hypothesis 6</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td>Hypothesis 7</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>Hypothesis 8</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>Hypothesis 9</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>Hypothesis 10</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Hypothesis 11</td>
<td>101</td>
</tr>
<tr>
<td></td>
<td>The Pupil Evaluations Forms' Reliabilities</td>
<td>101</td>
</tr>
<tr>
<td>V</td>
<td>ANALYSIS OF THE RESEARCH'S DATA</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>Pupils' Conception of Their Learning</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>Pupil Evaluation as Feedback</td>
<td>108</td>
</tr>
<tr>
<td></td>
<td>Pupil Evaluation and Dissonance</td>
<td>111</td>
</tr>
<tr>
<td></td>
<td>Pupil Evaluations and Subject Matter</td>
<td>112</td>
</tr>
<tr>
<td></td>
<td>Pupil Evaluation and Its Format</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td>An Over-All Analysis of Pupil Evaluation</td>
<td>117</td>
</tr>
<tr>
<td>VI</td>
<td>SUMMARY, CONCLUSIONS AND RECOMMENDATIONS</td>
<td>121</td>
</tr>
<tr>
<td></td>
<td>Summary of Procedure</td>
<td>121</td>
</tr>
<tr>
<td></td>
<td>The Statistical Treatment</td>
<td>122</td>
</tr>
<tr>
<td></td>
<td>Findings</td>
<td>123</td>
</tr>
<tr>
<td></td>
<td>Conclusions</td>
<td>126</td>
</tr>
<tr>
<td></td>
<td>Recommendations for Further Study</td>
<td>129</td>
</tr>
<tr>
<td></td>
<td>APPENDIX</td>
<td>131</td>
</tr>
<tr>
<td></td>
<td>BIBLIOGRAPHY</td>
<td>141</td>
</tr>
<tr>
<td>Table</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>1</td>
<td>Analysis of the Number of Responses Pupils Listed that They Learned</td>
<td>67</td>
</tr>
<tr>
<td>2</td>
<td>Percentages of Pupils' Responses Classified According to Eight Categories</td>
<td>71</td>
</tr>
<tr>
<td>3</td>
<td>The Total Number of Changes Each Student Teacher Made as Perceived by His Pupils</td>
<td>77</td>
</tr>
<tr>
<td>4</td>
<td>The Number of Pupil Groups Who Showed Change on Each of the Pupil Evaluations</td>
<td>79</td>
</tr>
<tr>
<td>5</td>
<td>Item Significance Summary for Pupils of Student Teachers Who Took an Initial Self-Evaluation and Those Who did Not</td>
<td>81</td>
</tr>
<tr>
<td>6</td>
<td>The Summary of Item Significance of Pupils Taking only a Final Evaluation and Those Taking Both Initial and Final Evaluations</td>
<td>84</td>
</tr>
<tr>
<td>7</td>
<td>Indications of Significance between Ratings by Pupils in Elective and Required Subjects for Form I</td>
<td>86</td>
</tr>
<tr>
<td>8</td>
<td>Indications of Significance Between Ratings by Pupils in Elective and Required Subjects for Form II</td>
<td>88</td>
</tr>
<tr>
<td>9</td>
<td>Item Significances Between Initial Student Teacher Self-Evaluations and Pupil Evaluations</td>
<td>90</td>
</tr>
<tr>
<td>10</td>
<td>Item Significance Between Final Student Teacher Self-Evaluations and Final Pupil Evaluations</td>
<td>92</td>
</tr>
<tr>
<td>11</td>
<td>Summary of Item Significances Between Initial and Final Student Teacher Self-Evaluations</td>
<td>94</td>
</tr>
<tr>
<td>12</td>
<td>Group Changes for Twenty-Six Student Teachers Between Experimental Groups' Initial Evaluations and Control Groups' Final Evaluations</td>
<td>96</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>13</td>
<td>Group Changes for Twenty-Six Student Teachers Between Experimental Groups'</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td>Final Evaluation and Control Groups' Final Evaluation</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Intercorrelations of Items on Pupil Evaluation Forms</td>
<td>99</td>
</tr>
<tr>
<td>15</td>
<td>The Pupil Evaluation Forms' Reliabilities on Each Item</td>
<td>103</td>
</tr>
<tr>
<td>16</td>
<td>Responses Made by Speech and Drama Pupils on Form I</td>
<td>132</td>
</tr>
<tr>
<td>17</td>
<td>Responses Made by English Pupils on Form I</td>
<td>133</td>
</tr>
<tr>
<td>18</td>
<td>Responses Made by Speech and Drama Pupils on Form II</td>
<td>134</td>
</tr>
<tr>
<td>19</td>
<td>Responses Made by English Pupils on Form II</td>
<td>135</td>
</tr>
</tbody>
</table>
## LIST OF ILLUSTRATIONS

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A Model of Communication</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>A Possible Classroom Communication Situation</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>A Model of Shannon's Concept of Communication</td>
<td>27</td>
</tr>
<tr>
<td>4</td>
<td>Knower's Model of Communication</td>
<td>29</td>
</tr>
</tbody>
</table>
CHAPTER I

BACKGROUND OF THE PROBLEM

One of the major distinguishing characteristics of man is his ability to use symbols. For this reason the philosopher, Kenneth Burke, describes man as a symbol creating, symbol using, and symbol abusing creature.\(^1\)

The manner in which man uses symbols can be viewed in terms of communication theory. Essentially, man transmits symbols, verbal or non-verbal, over some channel to another person who receives these symbols and interprets them. In such a system, meaning ascribed to symbols by either the sender or the receiver may be different. There is no guaranteed one-to-one correspondence between the symbols sent by the communicator and those received by the listener. Since like symbols may invoke entirely different meanings in any communication between individuals, there must be some means whereby the receiver of a message may check with the sender as to the similar use of the symbols transmitted. In simple

terms such a checking system is referred to as feedback. Feedback is exceptionally important in any communication system. Without it, much of the efficiency resulting from the use of symbols is lost.

All the current models of communication seem in agreement about the necessity of feedback in a communication system. Consequently, analysis of any communication system must give considerable attention to how feedback can maintain equilibrium within the system. If the sender's use of symbols is not understood by the receiver, or if the receiver is not free to give feedback to the sender, then there is little hope of maintaining for long, a workable communication system.

There exists a communication system which is not often viewed as such. More often it is discussed in terms of methods, processes, and achievements. The system being referred to is more commonly designated as teaching.

For the moment, let us consider the elements within the teaching-learning environment which cause the author to view it as a communication system. Considering the current models of the communication cycle, the essential elements of this system are a communicator, who sends a message over a channel, using symbols,

\[\text{In Chapter II a discussion will be presented on the representative theories of communication.}\]
to a receiver, who has the capability of interpreting these symbols, and upon interpretation, responds to the communicator with feedback. Figure 1 presents a simplified model of the communication cycle.

Fig. 1— A Model of Communication

Primarily the teacher is the communicator in a classroom for he has a message, the subject matter, to transmit to the pupils. The symbols most often used are verbal and commonly called language. But this does not mean there are not any use of non-verbal symbols such as facial expressions and bodily movement, gestures, stance and other possible aggressive or apathetic behaviors. As a receiver, the pupil interprets the teacher's symbols and thus tries to gain some understanding of the message or subject matter. At times feedback is returned to the teacher in the form of interpretable facial expressions, responses to questions asked by the teacher, questions of clarification of the message are asked by the students, or delayed feedback is obtained from examinations. Figure 2 depicts what a typical
communication system in a classroom may look like.

Fig. 2. — A Possible Classroom Communication Situation

Student₁ has received the message from the teacher and responds with interpretable facial expressions. Student₂ interpreted the message and became curious about it. First he asks Student₁ about what the teacher has just said and then because Student₁ wanted to continue to listen to the teacher undisturbed, Student₂ had to raise his hand and interrupt the teacher with his question. Student₃ received the message and thought it was important enough to tell Student₅ to pay attention. Student₅ replied to Student₃ by telling him to mind his own business. Student₄ was sleeping, so he missed all the communication that had taken place. As can be seen, there may be many different possibilities of where the communication cycle begins and ends.

If we evaluate the effectiveness of the communication system of a classroom, we in effect are analyzing the effectiveness of teaching and the learning consequently incurred. To study
communication in the classroom, some means of measuring communication behaviors must be employed. In other words, an instrument of feedback is necessary. One of the purposes of this study was to develop such a form which would validly and reliably measure communication in a classroom.

Statement of the Problem

The primary question to be investigated is, "Will early pupil evaluation of student teachers make more effective communicators of those student teachers, as perceived by the pupils?" The pupil evaluations are to be viewed as a form of feedback which may or may not have an affect on the behavior of the student teachers. Improvement will be measured in terms of the behaviors presented on the evaluations to be developed for this study. Thus higher numerical ratings on a second evaluation would connotate improvement as perceived by the pupils.

One of the important premises of this study was that student teachers need to become aware of the viewpoints of pupils in the learning situation. Student teachers need to discover whether a communications gap exists between them and their pupils. As novices, these teachers need to know from those to whom their symbols are directed whether or not they are "getting through."

Gage, Leavitt, and Stone contend that pupil evaluations are highly reliable, "pupils have greater opportunity than anyone
else to observe the teacher, and they are significant in number to make their pooled ratings highly reliable. Similar findings were reported by Howsam who had surveyed the research on pupil evaluations.

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

Since teaching is an act of communication, it needs feedback from the receivers of the messages in order to be effective. The instrument formulated for the elicitation of feedback from pupils focused upon the elements of communication which cause learning rather than those which cause liking or disliking of the teacher's behavior. The criteria evaluated by the pupils can be perceived as principally related

(a) to the teacher as a sender of messages,

(b) to how pupils receive the symbols transmitted, and

(c) to communication behaviors that could be observed by the pupils and altered or changed by the teacher.

---


In past studies, feedback was often obtained from pupil achievement examinations. These exams attempted to discover how much the student had learned. However, these examinations do not lend themselves to analysis of the communication system upon whose effectiveness they assume. If a teacher can know more specifically why the students did not learn the subject matter, he may be able to alter his communication behavior and yield better student learning and understanding. Often these written examinations measure such pupil attributes as memory, note-taking ability, and reading ability.

Another rationale often used in teacher evaluation is to have pupils rate their teachers on behaviors that supposedly are possessed by "good" teachers. But research has yet to state with confidence what these traits are. Strom and Galloway contend that "The traditional aim of trying to identify 'good' teachers and 'good' teaching has failed. Essentially it has failed because a single set of criteria has been used that defines success in an impersonal way." 5

From the many studies conducted at the University of Wisconsin under the directorship of A.S. Barr on teacher effectiveness, it has been concluded,

At this point [1961] it is emphasized that the

talents of individuals are highly diversified and that it is
doubtlessly psychologically unsound to expect many high level
talents in any particular individual or that they will be
highly intercorrelated.... In light of current theories of
the structure of human abilities, it is probably best to
hypothesize that teaching ability is composed of some parti-
cular combinations of special and general abilities. Much
of the research relative to the evaluation and prediction of
teacher effectiveness does not seem to reflect this point of
view to any considerable extent.5

Gage, Runkel and Chatterjee developed a pupil feedback
form based upon the assumptions of equilibrium theory. Pupils
rated their teachers upon twelve behaviors and they then gave
ratings on the same criteria for their "ideal" teacher. The
results of the study were inconclusive.

Teachers did change, in the direction of pupils' ideals, as
described by pupils, as a result of getting feedback. Whether
the changes were great enough to have educational significance,
whether they would be found if teachers' behaviors were de-
scribed and measured by expert outside observers, rather than
pupils, whether the changes toward educators' ideal — all of
these are questions for subsequent investigations.7

Considering these research findings and others to be
discussed in Chapter II, it has been hypothesized that a new
approach to the concept of evaluation is needed. It seems that

6A.S. Barr, et. al., Wisconsin Studies of the Measurement
and Prediction of Teacher Effectiveness,(Madison: Dembar Pub.,

7N.L. Gage, P.J. Runkel, and B.B. Chatterjee, "Equilibrium
Theory of Behavior Change: An Experiment in Feedback From Pupils
to Teachers," Report No. 6 in Studies in the Generality and
Behavioral Correlates of Social Perception, (Urbana: Bureau of
Educational Research, College of Education, University of Illinois,
1960) p. 94.
pupil evaluation has been viewed as nothing other than rating or grading of teacher. But teaching is an interpersonal act of communication in which there is an attempt to influence others. ¹⁸

"Any classroom is a small communications world in itself. The key person controlling classroom communication is the teacher."²⁹

For these reasons pupil evaluation or feedback must be looked upon as a more intimate and personal endeavor. The student teacher must personally solicit feedback from his pupils. The pupils must feel that they are cooperating in improving the learning situation by responding on an evaluation form. Psychologically, the student may feel that he has more of a stake in making the learning environment worthwhile if his opinions are sought.

Just by asking the pupils to evaluate the communication system within the classroom, a student teacher may improve the pupils' reception of his messages. Rather than always telling pupils how to behave, the teacher, by giving to students some

¹⁸S. Lundstedt, "Criteria for Effective Teaching," Improving College and University Teaching, XIV(Winter, 1966), p. 27.

form to evaluate him, helps complete the communication cycle by eliciting feedback. To be effective, communication cannot be one-sided; it must maintain interaction between the communicator and the receivers.

In using the term feedback, it must be understood that some modification of communication behavior occurs upon the receipt of feedback. The communicator may continue, discontinue, or adopt a new behavior or distort the feedback and dismiss it as valueless. If there is no opportunity to respond upon the reception of feedback then there seems to be limited value of procuring feedback. Most studies in the past have obtained evaluations of student teachers at the end of their experiences. These studies seem to have taken a very narrow view of evaluation. A basic assumption upon which final evaluations rest is that a student teacher will have classes in the future similar to the one doing the evaluation. In some instances this may be true, but the contrary is far more probable. Very few groups of pupils will ever behave similarly under the same teacher at a different time. The teacher and the students will change. Evaluation should not be a one-shot endeavor but should be a continuous process operative throughout the entire program of student teaching as well as continued into regular teaching. "Sound evaluation is developmental and endeavors to appraise growth throughout the entire period of teaching."\(^{10}\) Hopefully if a student teacher asks pupils to

evaluate the communication system before the end of their association, he may be able to make adjustments and changes which will enhance the learning among his pupils.

Only a few studies have used pupil evaluations of student teachers prior to the end of their experience. Using the Purdue Rating Scale for Instructors, Remmers, Ward and Schmalzried found "As a result of the training program, thirty-nine (out of 40) of the practice teachers made an average gain in rating of 4.2 scale points, which is educationally and socially significant as well as statistically significant."[11]

In the Remmer's study the first pupil evaluation was given after the student teachers had taught their classes for four weeks. After an interval of 12-14 weeks, the end of the semester, the second evaluation was made by the pupils.

In the dissertation study of Savage[12] the first evaluation was given only five days after the student teachers had begun to teach. The interval between the first and second evaluations


averaged about twenty days. Her results were not statistically significant, and even the trend was not in the hypothesized direction.

Taking into consideration these findings, it would appear that the proper time of evaluation is quite important. Not only must consideration be given to how early students can evaluate student teachers, but how long an interval is needed to obtain significant change. Since many student teaching programs are shorter than a full sixteen week semester, it is important to discover how many days pupils must be exposed to a student teacher before they can evaluate him and overcome any strong first impressions.

From the results of a pilot study and Savage's study, it appears that five days after the student teacher has begun teaching is too early to have students evaluate him. But if we wait for four weeks, as done in the Remmers, et al., study, there may be left too short an interval between the first and second evaluation for significant behavior changes upon the part of the student teachers. On a quarter system many student teachers do not start teaching for about two weeks, and giving an evaluation four weeks from this time would only leave two or three weeks between evaluations. It may take more time to affect significant
communication behavioral changes. Consequently this study is investigating the validity and reliability of gaining feedback from pupils after the student teachers have taught approximately ten days and then obtaining a second and final evaluation from four to six weeks later.

Feedback and Learning Theory

Another manner in which to view the feedback process is in terms of the general learning cycle. According to many learning psychologists, in order for learning to occur, three conditions must be present. There must be a motive present which seeks to obtain some goal, but unless there is a block or barrier making this goal not immediately attainable, learning will not occur. "This general pattern dominates the whole human behavior -- from birth to death."\(^{13}\)

Evidently the key to the learning cycle is the presence of a block. Therefore, if we desire the student teacher to learn during his field experience, we must establish a block toward the gaining of his goal, success in communicating with his pupils. Feedback, or pupil evaluation, may serve as this block if there is a means of determining the student teacher's self-concept. In other words, a student teacher's self-evaluation may be compared with the average pupil evaluation.

On criteria where there is significant disagreement, and there should be at least a few, according to equilibrium theorists, dissonance occurs.

**Feedback and Equilibrium Theory**

Since humans usually seek to eliminate dissonance, a student teacher will probably try to discover the reasons why his pupils have different perceptions of his abilities as a teacher-communicator. In this search to overcome the block, the student teacher may learn how to improve his effectiveness in certain types of communication.

In support of student teacher self-evaluation Dumas states, "If a student teaching experience should facilitate the improvement of one's self-concept, we may assume that the experience has, to that extent enhanced the potential of that student teacher for actual success in teaching." Strom and Galloway add to the value of student teacher self-evaluation the dimension of professional growth. "In essence, growing acceptance has accrued to the position that continued competence based upon self-insight into one's own behavior and its consequences

---

is the major means of accomplishing professional growth."^{15}

The Instrument of Feedback

In order to evaluate the pupils' abilities to interpret the messages transmitted by the student teacher, the feedback instrument must somehow elicit an interpretation of the messages. Rather than obtaining a measure of learning by inference, i.e., as that obtained from written tests, we may be more direct and ask the pupils what they have learned or not learned. In no study investigated evaluating the effectiveness of teaching has there been such an open query directed towards pupils. From an analysis of the pupil responses, we may discover whether pupils are aware of learning anything other than facts, which require the least complex of the thought processes, memory. Considering Bloom's Taxonomy^{16}, we need to determine if pupils are aware of learning more complex forms of thought such as translation, interpretation, application, analysis, synthesis, or evaluation.

\[^{15}\text{Strom and Galloway, op. cit., p. 287.}\]

\[^{16}\text{Benjamin S. Bloom (ed.), Taxonomy of Educational Objectives, New York: Longmans, Green, 1956.}\]
Questions and Hypotheses to be Studied

Some of the questions for which this study hopes to provide some tentative answers are:

1. How soon can pupils give reliable and valid feedback of the communication behaviors of student teachers?

2. Do the evaluation forms developed for this study present an adequate picture of the total teaching situation?

3. Will pupil feedback affect the communication behaviors of student teachers as perceived by the pupils?

4. Do student teacher self-evaluations appear realistic on either early or final administrations?

5. Do pupils evaluate their student teachers with any form of a halo effect?

6. Are pupils able to itemize types of thought processes taught which emphasize more than memory learning?

In order to answer these questions, most of them have been incorporated into a set of null hypotheses. To facilitate reference to any of these questions in the ensuing manuscript, they will be coded Q1 through Q6.

The null hypotheses to be accepted or rejected by statistical analysis are the following:

H1. There is no significant difference in pupil evaluation of student teachers between an early and a final evaluation.

H2. There is no significant difference in final pupil
evaluations between student teachers who did and did not take the initial self-evaluations.

H3. There is no significant difference between pupil evaluations of student teachers who were evaluated initially and those who were given only a final evaluation.

H4. There is no significant difference between pupil evaluations of student teachers in required subjects and of those in elective subjects.

H5. There is no significant difference between a student teacher's initial self-evaluation and the average initial pupil evaluation of communication behaviors.

H6. There is no significant difference between a student teacher's final self-evaluation and the average final pupil evaluation of communication behaviors.

H7. There is no significant difference between student teacher's first and final self-evaluations.

H8. There is no significant difference between initial pupil evaluations of a student teacher and final evaluations done by pupils who did not take the initial evaluations.

H9. There is no significant difference between final pupil evaluations of a student teacher and those in the control class.

H10. There are no significant differences among intercorrelations of the items on Form I.

H11. There are no significant differences among intercorrelations on Form II.
For future cross-reference the above hypotheses will be coded HI through HU.

**Design of the Study**

In order to obtain data, this study was conducted during the winter and spring quarters of 1969 at The Ohio State University. Student teachers in the fields of Speech, English, History, and Science were used as subjects. They taught in junior and senior high schools in the Columbus, Ohio Public Schools and surrounding school districts.

Two parallel pupil evaluation forms were developed for this study, copies of which will appear in the appendix. Each form asked the pupils to evaluate their student teacher upon twenty-five communication behaviors and one general effectiveness item. The second half of the form sought to elicit from the pupils what they had learned or not learned from the student teacher. In order to compensate for memory factor, the wording on each form was slightly varied and the sequence of items was reversed. Pupils were asked to evaluate on the quantity of occurrences of various communication behaviors rather than on their quality. A seven-point continuum from always - never was used instead of one connotating the idea of grading such as a continuum using excellent to poor. Hopefully a quantifying scale would reduce the tendency among pupils to feel that they...
are assigning a grade to the student teacher. All but one item was stated in the positive sense, that is, if an item is given a value of 7 (always), we should conclude that this communication behavior was operating and effective in the eyes of the pupils. The single item stated in the opposite or negative sense was inserted in the series of statements as a benchmark item. If a pupil was not reading the statements carefully and just marking high or low scale values on all the items, he would do the same for this differently stated item, thus the evaluation could be rejected as invalid.

The design of the study was made as simple as possible so as to not introduce any extraneous variables that might distort the results. During the winter quarter, 1969, student teachers in Speech, English, Science and Social Studies were asked to administer the pupil feedback forms to one of the classes that they were teaching after the tenth day of teaching. Prior to administering this form to their pupils, some of the student teachers filled out a self-evaluation on the same communication behaviors. Upon analysis of the pupil evaluation forms, the student teacher could, if he wished, discuss with the cooperating teacher and the college supervisor the areas of weakness and strength in the communication system of the classroom as depicted by the pupils. At the end of the field experience, these same student teachers administered a parallel feedback form to their pupils.
To gain further data, the study was replicated during the Spring quarter, 1969. Only student teachers in Speech and English participated in this part of the study. The same procedures were followed for the pre-and post-test conditions.

**Limitations of the Study**

A study of the pre-described nature is certainly not without limitations. The author wishes to openly state some of the major limitations which may directly affect the generalizations made from an analysis of the data collected.

1. The study cannot claim that the criteria upon which student teachers were being evaluated were the only and most important ones. Moreover, it was hypothesized that they were a representative sample.

2. Control of the administration of the pupil evaluations by the student teachers was limited to their ability to follow directions and take the feedback process as a necessary part of professional growth.

3. Change in communication effectiveness is restricted to the feedback form's format and ability of the pupils to discriminate changes in the student teacher's communication behaviors over a period of weeks.

4. There were some inaccuracies in interpreting items by
both pupils and student teachers because of the variableness in the meaning of words.

**Definition of Terms**

There will be a number of terms used throughout this manuscript that are given a specialized meaning. To avoid possible confusion, the author will list these words with his description of their usage or meaning.

**Feedback** -- Student responses on the communication behaviors of student teachers. It will be used interchangeably with the terms pupil evaluation, pupil reaction, or pupil rating.

**Communication behaviors** -- They are observable characteristics of any element of the communication cycle which are alterable.

**Self-evaluation** -- It is the student teacher's analysis of his own communication abilities in a classroom.

**Significance of the Study**

Now that some definition and design of the present study has been presented, the author wishes to consider the significance of the research conducted.

Educators are constantly saying that the pre-service education of teachers needs improvement. Since student teaching is an important part of this pre-service program, it is paramount that the experience be a worthwhile learning endeavor. But little
has been done to improve the field experience besides altering its length or changing its sequence in the future teacher's course of studies. An important aspect of any communication experience, student teaching being such an experience, is to elicit feedback from the receivers of the messages sent by the communicator who, in this case, is the student teacher.

Since there are many conflicting results in the studies of teacher evaluation, there must be continuous investigation of the problem until some reliable trend is established. This study hopes to contribute to the establishment of the trend toward more general acceptance of the validity and reliability of pupil evaluation of teachers.

For the first time, this study combines the theories of communication, equilibrium and learning, to investigate the still vaguely defined phenomena of teaching. The study elevates the process of pupil evaluation to a cooperative effort to improve the learning situation rather than viewing it only as a rating or grading of the student teacher's efforts.

The Plan of the Dissertation

To further explicate the reading of the rest of this manuscript a summary of the first chapter and a preview of the content of the following chapters seems necessary.

The first chapter included the background of the issues, the definition of the problem, the hypotheses to be tested,
the design of the study, the limitations of the study, the definitions of key terms, and the significance of the study.

Chapter II will contain a discussion of the related theories of communication, learning and equilibrium which contribute to the development of the rationale of the present study. Also to be found in this chapter will be a review of the literature on pupil evaluation of regular and student teachers.

In Chapter III, a discussion of the sample used, methodology employed and the development of the feedback form will be presented.

Since the pupil evaluation form solicits two kinds of information, Chapter IV will contain the statistical analysis of both types of information. Analysis of what pupils perceive they are learning will be presented along with the data related to accepting or rejecting the null hypotheses. Interpretation of the statistical data presented in the previous chapter will be found in Chapter V.

The concluding chapter, Chapter VI, will present the summary, conclusions drawn and recommendations suggested from this study.
CHAPTER II

REVIEW OF THE THEORY AND LITERATURE

To develop the rationale for this study, it was necessary to investigate theoretical concepts and previous experimentation in the field of pupil evaluation of teachers. Ideas presented are those which are believed to be contributive to the development of a unified theory. The specific theories that will be considered and integrated are the theories of communication, learning, and equilibrium. From the integration of these theories, the basis of the present research may become more evident.

Theories of Communication

Ever since man began using symbols, he probably has been trying to develop a comprehensive theory of communication. The first recorded attempts at developing such a theory can be found in the writings of such Greek scholars as Isocrates, Demosthenes, and Aristotle. It was Aristotle who composed the first comprehensive treatise on communication. This systemization was developed in his book on rhetoric. The orientation or primary emphasis of Aristotle's theory of communication was on the message. Though Aristotle presented a very detailed elaboration on the preparation,
characteristics, and methods of the speaker, he failed to develop a complete theory of communication.

For two millenia no significantly different theory of communication was originated. Three British rhetoricians, Blair, Campbell, and Whately (who lived in the eighteenth century) recognized the incompleteness of Aristotelian rhetoric or theory of communication, and advanced a new approach to the viewing of communication. Their orientation was psychological and emphasized the speaker-listener relationship.¹

Not until the present century has there developed a renewed interest in creating a comprehensive theory of communication. Two of the leading scholars who have been instrumental in the development of a fresh and more comprehensive view of communication are I.A. Richards and Kenneth Burke. Concerning this new approach to the study of communication or rhetoric Richards says,

In rhetoric's care is that unity, or, as Coleridge would have written, coadunation of the mind, which is, whatever the deviations, the aim behind and before our strivings -- a necessary organic interanimation of meanings, the biological growth of the mind in the individual and in a social inheritance maintaining human advance.²

Burke holds that all the new disciplines have contributed


to the development of contemporary rhetoric. Within the scope of rhetoric Burke places, "all those statements by anthropologists, ethnologists, individual and social psychologists, and the like, that bear upon the persuasive aspects of language the function of language as addressed, as direct or round-about appeal to real or ideal audiences, without or within."\(^3\)

The sociological orientation of these two scholars has greatly influenced modern theories of rhetoric or communication. By sociological is meant that communication theory is concerned with more than separate entities such as the speaker, the message, or the listener. A sociological orientation approaches the study of communication by a study of the interrelationships of each element in the communication cycle.

Another field that has had its affect on the development of a modern theory of communication is Cybernetics. The important concept of cybernetics that has relevance to communication is feedback. In cybernetic theory feedback is referred to as a servo-mechanism which automatically adjusts its output according to feedback signals that supply information about how closely the controlled action is approximating the desired condition.

From the concept of cybernetics has evolved information theory. Shannon was instumental in proposing such a theory.

\(^3\)Burke, op.cit., pp. 43-44.
Essentially Shannon's theory deals with a source selecting a message, encoding of this message by a transmitter, sending it over a communications channel which may possess some noise or interference, and then decoding of the message by a receiver at a destination.4

![Diagram of Communication Model]

Fig. 3. — A Model of Shannon's Concept of Communication

Following Shannon's innovation were a number of theorists who developed models of communication. Lasswell proposed that communication theory and research are concerned with, "Who says what through what channels to whom with what effect."5

Gerbner's verbal model was somewhat more detailed. In the following terms he identified ten aspects of communication, "(1) Someone (2) perceives an event (3) and reacts (4) in a situation

---


(5) through some means (6) to make available materials (7) in some form (8) and context (9) conveying content (10) of some consequence."

This search for an adequate model or theory of communication has given rise to a new field of study, designated as communicology. A major contributor to this new field has been Franklin H. Knowler. He has developed a comprehensive model of communication which combines all the elements of the previous theories. Essential to Knowler's theory is the adapting of a message for a receiver on the basis that all communication behavior is learned and is subject to modification through continued learning.7

Another communication theorist Berlo, even goes so far as to suggest that in analyzing the process of learning, we in effect, are analyzing the process of communication. He states,

The six ingredients that are involved in learning have their analogies in the ingredients that are involved in communication.

<table>
<thead>
<tr>
<th>Ingredients in Learning</th>
<th>Ingredients in Communication</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Organism</td>
<td>1. Channel</td>
</tr>
<tr>
<td>2. Stimulus</td>
<td>2. Message</td>
</tr>
</tbody>
</table>


7Franklin H. Knowler, "A Model for a Communicology," Ohio Speech Journal, II (Spring, 1963), p. 183. Dr. Knowler has given his kind permission to allow the presentation of the model on the following page.
Fig. 4 — Knower's Model of Communication
3. Perception of Stimulus  3. Decoder  
4. Interpretation of Stimulus  4. Receiver-Source  
5. Overt Response to Stimulus  5. Encoder  

Subsequently, if teaching is viewed as an act of communication, then the teacher behaviors can be learned and modified with the aid of feedback from the receivers of the messages, the pupils.

In order to see more definitely how closely related are the theories of communication and learning, it seems worthwhile to review two major contemporary theories of learning. From this analysis, the author hopes to make more explicit the basic assumptions of this study.

Theories of Learning

Prior to a discussion of any theory of learning, it is pertinent to take cognizance of Hilgard's admonition, made in 1956, which still holds true,

There are no laws of learning which can be taught with confidence. Even the most obvious facts of improvement with practice and the regulation of learning under reward and punishment are matters of theoretical dispute.°

The theories to be discussed appear to support the rationale


of the author's investigation. Understanding these theories will facilitate explanation of phenomena being studied.

In 1943, Hull published the very influential volume, *Principles of Behavior*. A theory, designated as reinforcement theory, was developed. The condition of need in organisms is an important factor in Hull's theory. Need "plays an important role in determining the occasions when habits shall function in the evocations of action, the vigor of such evocations, and their persistence in the absence of reinforcement."10 Such a position can be assimilated into an information or communication theory. "The minimum reinforcement necessary for habit strength is that required for information for feedback; excesses beyond that effect reaction potential but not habit."11 To summarize Hull's theory: learning requires the occurrence of reinforcement, and that temporal contiguity in of itself is not enough to produce learning.

Mowrer offered a revision of Hull's theory. He felt that much emotional learning, such as conditioned fear, is a result of simple contiguity conditioning without necessarily involving reinforcement through drive reduction. In a later revision of his own theory, Mowrer proposed that all learning is


sign learning or conditioning by contiguity, and instrumental learning is just a special case. "Sign learning is often characterized as involving stimulus substitution or associative shifting, whereas solution learning is often described as involving response substitution."\textsuperscript{12}

Considering the above theories of learning and communication, the author perceives that they can be related or integrated into theories of equilibrium.

\textbf{Theories of Equilibrium}

To review what the goals of perceptual learning are, Hilgard pointed out the necessity of the individual to achieve environmental stability.

The organism seeks a perceptually stable environment in a fashion somewhat parallel to that in which it seeks an internally stable environment. There is a kind of environmental homeostasis parallel to physiological homeostasis. The equilibrium is a dynamic one, and the external environment, like the internal one, is ever changing.\textsuperscript{13}

Festinger deals with the same ideas in his theory of cognitive dissonance. Dissonance occurs when there is disequilibrium or imbalance in the individual. "Cognitive dissonance can be seen as an antecedent condition which leads to activity oriented toward dissonance reduction just as hunger leads to

\begin{footnotesize}
\begin{itemize}
\item[\textsuperscript{13}] Holgard, \textit{op. cit.}, p. 466.
\end{itemize}
\end{footnotesize}
activity oriented toward hunger reduction.\textsuperscript{14}

According to Festinger, the presence of dissonance leads an individual to seek information (feedback) which will help him restore cognitive balance. When dissonance occurs as a result of others' reactions to him, a person will solicit these individuals' opinions or reasons for their imblance-causing reactions. Consequently, how much one values the opinions of those causing dissonance is directly related to its reduction after feedback.

The magnitude of the dissonance introduced by disagreement from others increases with the increase in the importance of the opinion to the person, in the relevance of the opinions to those voicing disagreement, and in the attractiveness of those voicing disagreement.\textsuperscript{15}

If an individual sees little value in soliciting the opinions of a group which causes dissonance because of their disagreement of his self-concept, then obtaining their opinions is worthless. On the other hand, if he sees the values of receiving the opinions of many with whom he is associating, he will more easily reduce dissonance and even take precautionary actions to avoid situations which may cause dissonance. In situations of interpersonal communication, it is extremely important to maintain an open and balanced system. But when dissonance occurs in an interpersonal communication system, it can more easily and


\textsuperscript{15}\textit{Ibid.}, p. 263.
quickly be eliminated if there is honest and meaningful feedback. Sometimes this conceptual conflict between persons will not emerge immediately, but wait until some external stimulus pattern, verbal or non-verbal, causes incompatible symbolic responses to be aroused in combination.\footnote{16}

Having considered all the above theories, we are now ready to integrate or inter-relate them to the present study.

Integration of Theories

Student teachers are not initially able to recognize possible conceptual conflicts between themselves and their pupils. At first, they are engulfed in the complexity of the task of teaching and have difficulty in gauging their ability to communicate with their pupils. Since about 80\% of their time is spent in communication, it is important that they discover whether there is disagreement in their own conception of their ability to communicate to their pupils and the pupils' reactions to their communication behaviors. If a student teacher obtains, from pupils, areas of disagreement, he may, according to dissonance theory, change his own opinions of his communication abilities. Two other avenues of action are also open: the student teacher may try to influence the pupils to change their opinions, or he may just reject pupils' opinions as being invalid.\footnote{17}


\footnote{17}{Festinger, \textit{op.cit.}, p. 265.}
Using either of the first two alternatives would improve the communication system between the student teacher and the pupils because either action causes a re-evaluation of the messages sent and those received. The third alternative of rejecting the pupils' opinions entirely is in direct opposition to the learning process.

In this study the student teacher is viewed primarily as a communicator. The pupils are the receivers of the messages he transmits. Therefore, if a student teacher desires to obtain feedback of his communication ability, it would seem equally logical to solicit the opinions of the receivers or pupils, as well as the opinions of the observers, the cooperating teacher or the college supervisor. In order to determine whether there are incompatible opinions on communication behaviors between the student teacher and his pupils, it is important to establish a format from which to exchange feedback from the pupils.

Another factor that needs consideration is the time at which pupils are asked for their opinions. If one is to wait until the end of the student teaching experience, the student teacher may discover that he did or did not communicate well, but there would be no way of reducing the dissonance that may occur for the student teacher will no longer be associated with the class. If a student teacher obtained evaluations of his communication behaviors early in the experience, there is the possibility that any resulting dissonance may be reduced in the period of time...
following the evaluation. Because of the nature of communication, it would be highly improbable that the average level of agreement on communication behaviors between a group of pupils and their student teacher would be perfect. If the pupil feedback format was so structured, it would present upon analysis, a profile of the strengths and weaknesses of the student teacher's communication ability as viewed by his pupils. Communication behaviors which the student teacher is proficient in would be reinforced and ones of lesser competence may cause dissonance. If dissonance is aroused on certain communication abilities, then given some time, the student teacher may endeavor to improve or may reject the pupils' opinions and possibly become even weaker. One way to determine the occurrence of such possibilities, after some interval has intervened, is to ask the pupils to re-evaluate the student teacher's communication abilities. A comparison of the average pupil response on each identical trait may signify improvement, stability, or weakening. Thus it was the purpose of this investigation to have student teachers' communication abilities evaluated by their pupils near the beginning and at the end of the field experience. Not only may the student teacher become more quickly aware of the pupils' perceptions of his communication prowess, but he may develop a better communication system because he is soliciting and using feedback.

Previous studies have also sought pupil evaluations of
teachers, but conclusions of their value have varied greatly. Probably the difference in formats and rationales have made it somewhat difficult to relate the conclusions reached by different investigators. From a study of the literature one may come to understand the value and need for further study of pupil evaluations of teachers.

To find consensus on the value of pupil evaluations of teachers is difficult. But even more difficult is the task of finding studies using pupil evaluation of student teachers. Thus much of the review of literature deals with pupil evaluation of regular teachers. It seems educationally sound to propose that results of studies with regular teachers can be generalized to student teachers since teaching is a developmental profession in which members are discovering new ways of becoming more effective teachers. Although student teaching is a shorter experience, the student teacher usually gains the full responsibility of a class or classes by the end of the term. He is expected to respond with behaviors appropriate to a regular teacher.

The researchers of pupil evaluations of teachers who have found some positive results in their studies are quite enthusiastic in their endorsement of the evaluations educational values. In the next section will be found some of their optimistic comments.
Values of Pupil Evaluations of Teachers

One of the most prolific researchers in the area of teacher evaluation, Roy C. Bryan, distinguishes between teacher evaluation and teacher ratings done by pupils. His interpretation of pupil evaluations has been one of the strong stimuli upon which the present study was conceived.

Pupil responses should not be regarded as ratings of the teacher. They are simply a report on pupils' opinions, feelings, and prejudices.... If they agree that they see no value in what is being taught, we must agree that they see no value in what is being taught. It does not necessarily follow that their opinions are right.  

Another persistent concern in teacher evaluation is that evaluations present a fair representation of the teacher's complex of responsibilities. Seldom can external observers, unless they continually stay in a class for a number of weeks, gain a full insight into reasons for a teacher's actions. Usually the observer does not know the pupils personally, nor is he aware of previous pupil-teacher interaction which may have precipitated the particular behaviors he is witnessing. Since outside observers visit a teacher sporadically, their presence may cause atypical pupil behavior as well as atypical teacher responses.

This line of reasoning contradicts the position of David G. Ryans, the author of Characteristics of Teachers.

---

According to Ryans, "only with training of observers can one expect to obtain meaningful assessments of teacher behavior; it is the only proper way one can approach teacher assessment for either research purposes or for pre-service or in-service teacher evaluation."\textsuperscript{19}

It seems in holding such a view, Ryans, denies the possibility that students can validly and reliably assess teaching behaviors. According to Ryans, only trained observers with pre-determined ideas of what is effective teaching can randomly observe teachers and evaluate them fairly. But as pointed out in Chapter I by Strom and Galloway,\textsuperscript{20} research has yet to establish an incontrovertible definition of good or effective teaching. A more realistic approach to evaluation of teachers would be to observe a teacher's performance in a wide variety of roles demanded of him. Here is where students have a major advantage over outside observers. "They see the teacher perform on many different occasions as he attempts quite varied tasks and deals with individuals who are known to the observer."\textsuperscript{21}

Since pupils know each other much more intimately than could any outside observer, they often know why a teacher is reacting


\textsuperscript{20}Supra, p. 7.

\textsuperscript{21}I. Paraskevopoulos, "How Students Rate Their Teachers; Rogge's Style of Teaching Inventory," Journal of Educational Research, LXII(September, 1968), p. 25.
to or dealing with a fellow-pupil. Thus their gaging of the fairness or appropriateness of the teacher's behavior may be very different from the impartial outside observer. For example, if a teacher "blows up" at a pupil, an observer may designate the behavior as an uncalled for burst of temper. Yet, pupils may view it as justified because they have seen this pupil goad and aggravate the teacher for the previous few days.

In addition to the important advantages accrued from many observations of teachers upon which to base judgments, Veldman and Peck point out, "the use of pupils as observers also affords the increased reliability and reduction of bias that multiple judges afford."22

An educator who conducted a study similar to the present one maintains,

A significant factor in the success of a student teacher is the kind of effect she has on pupils, or how pupils perceive her characteristics and performance. Sensitivity to pupil reaction early in student teaching is desirable if a teacher is to perform effectively.23

In considering the reliability and validity of pupil ratings of student teachers, Bowman concluded from his study,

1. The morale of the pupils was improved by the opportunity to rate their teachers.


2. Student teachers welcomed pupil evaluations as a means of growth.
3. The more often pupils were asked to evaluate teachers, the more understanding they appear to have and the more seriously they regard the undertaking.
4. When corrected for attenuation, such evaluations appear to be about as reliable as results on many standardized tests.
5. On the whole, pupils' ratings seem to be little affected by the marks they receive directly from the student teacher.

Students are capable of perceiving criteria necessary for good teaching that can be supported with sound educational theory. Bryan found in one of his studies that the following criteria usually topped the lists comprised by students:

1. Knowledge the pupils think the teacher has on the subject,
2. Ability of the teacher to explain things clearly,
3. Amount of work done by the teacher, and
4. How much the students think they are learning from the teacher.25

H.E. Cripe adds another value of pupil evaluations.

"Pupil evaluations, unlike other types, seem to be able to meet the major objective, namely, to stimulate the teacher toward self-improvement."26 With early, rather than just final, pupil evaluations, a teacher may have areas of weakness more clearly designated. Specific goals for improvement can more carefully and specifically be determined after a study of the pupil evaluations.

---


and often teachers may be able to eliminate unsuspected factors that are blocking effectiveness.  

Fetterhoff calls to our attention that student opinion surveys often inform the teacher of certain undesirable attitudes and features of instruction and personality of which he was not aware.

Blum concluded from his study that pupil evaluations "are as good indicators as is possible to obtain." He also found that pupil grades did not significantly effect the evaluations of the teacher's ability to teach.

Bryan asserted,

"When one thinks of student reactions as a means by which teachers can gain a better understanding of students and the effects of teaching on students, he has no justifiable reason to question validity.... Since students themselves are the primary and ultimate source of information on their opinions, we must accept their opinions as valid."

Guthrie found in his study that student opinion of teachers

---


30Roy C. Bryan, "Reliability, Validity, and Needfulness of Written Student Reactions to Teachers," Educational Administration and Supervision, XXVII (December, 1941), p. 656.
were highly reliable and that more consistency was found in student judgments of college faculty than faculty judgments of faculty.\textsuperscript{31}

In addition to the number of advantages already mentioned of pupil evaluations of teachers, Hood culled from an examination of over fifty reports some of the following:

1. Favorable findings of pupil evaluations may encourage teachers and build up their self-confidence; they may stimulate good teachers to become even better.
2. Teachers are made more aware of the great influence they have on their students and, as a result, may be motivated to greater care and effort.
3. Pupil evaluations of teachers have an element of privacy not found in ratings by principals, supervisors, or other teachers.
4. Sources of prejudice, anxiety, and worry of students may be traced to their reactions to teachers; pupil evaluations will enable teachers to work for a more wholesome atmosphere from the standpoint of mental health.
5. Democracy in the classroom is based on treating students like real people, so teachers should find out their pupils' feelings, opinions, and desires in order that they may respect them.\textsuperscript{32}

In a study of 1588 high school students evaluating their teachers, Albert presented the following values as derived from an analysis of his data:

1. Pupil evaluation of teachers is reliable, valid, practical


and inexpensive.

2. Teachers can be benefitted by pupil evaluations.

3. Pupils are sufficiently consistent in the evaluations of teachers for the results to be meaningful.\textsuperscript{33}

The values mentioned above have been found in studies of pupil evaluations of both regular and student teachers. Discussion of specific research though will be separated into a section on studies on regular teachers and then a section on studies conducted with student teachers.

\textbf{Studies in Teacher Evaluation}

In an extensive research project sponsored by the United States Office of Education, Bryan had as one of his objectives to determine to what extent may the feedback of information about student reactions improve teacher effectiveness. The group of teachers were all certified to teach. Two groups were formed, an experimental group of sixty teachers and a control group of fifty-nine teachers. The experimental group had from two to three of their classes evaluate their teaching during the spring of three consecutive years. Such pupil reactions were also received from the students of the control group. The control group never saw how their students evaluated them.

The ten questions upon which pupils were to evaluate their teachers using a four point continuum were:

1. The knowledge this teacher has of the subject matter?
2. The ability of this teacher to explain clearly?
3. The teacher's fairness in dealing with students?
4. The ability of this teacher to maintain good discipline?
5. The sympathetic understanding shown by this teacher?
6. The amount students are learning in this class?
7. The ability of this teacher to make classes lively and interesting?
8. The ability of this teacher to get things done in an efficient and business-like manner?
9. The skill of this teacher in getting students to think for themselves?
10. The general (all-around) teaching ability of this teacher?  

Bryan discussed some of the data in terms of gains or losses on one or more questions. Experimental teachers made more gains than the controls on all questions. From the tables presented it can be observed that 57% of the sixty teachers in the experimental group made significant gains on one or more of the questions as compared with 24% of the fifty-nine control teachers. One-sixth of the experimental teachers (10) made gains on six or more questions, while none of the control group did so.  

From the mass of data collected in this study, the author, Bryan, concluded that the data "indicate that the feedback of information about student reactions can be used by many teachers as a means of improving effectiveness as seen by students."

---

36 Ibid., p. 43.
Bryan's study dealt with regular teacher and final pupil evaluations. No discussion was presented on how much, if any at all, change occurred between the first and second evaluations. Therefore it is not possible to infer from the data that two years was a necessary time interval for the resulting changes. Since different classes of pupils evaluated the teacher each year, we cannot really determined if the makeup of the classes as groups of interacting individuals was responsible for differences in evaluations or it was teacher initiative. Possibly, if at the beginning of each year, these teachers were evaluated by their students, we could have determined if there was teacher behavior change. The experimental group need not have seen this initial yearly evaluation, but it may have been compared with the evaluations of the previous spring. If there were significant changes between these two evaluations, then, it would seem that teachers conscientiously analyzed their behaviors as perceived by the student in the previous year and were making efforts to cause changes in their behaviors. Such a conclusion cannot be made for Bryan's experimental conditions.

In another experiment conducted with certified teachers, improvement over a period of time with a particular class was studied. Gage, Runkle, and Chatterjee conceived their study in terms of equilibrium theory. The question their investigation sought to answer was, "Can teacher behavior be changed by informing the teachers
how their pupils describe the behavior of their actual and their ideal teacher? The pupils in the study were sixth graders in the state of Illinois. The teachers were randomly assigned to the experimental and control groups.

Assuming that the twelve behaviors evaluated are significant and observable by the students, this study was summarized in the following words: "Teachers did change in the direction of pupils' ideals as described by pupils, as a result of getting feedback. Whether the changes were great enough to have educational significance... [is a question] for subsequent investigation."^38

Realizing that teaching is more than a complex of twelve behaviors, one may contend one of the reasons for less than conclusive significance in change for the above study was that not enough behaviors were evaluated. It would be interesting to replicate this study and increase the number of behaviors evaluated by the pupils. Using one of the statistical prophecy formulas may have suggested the number of items needed for significantly higher reliability. It may be hypothesized that teacher changes in behavior may become more evident when a larger group of traits are evaluated.

Tuckman and Oliver conducted a study with 286 vocational

---

^37 Gage, Runkel, and Chatterjee, op.cit., p. 5.

^38 Ibid., p. 94.
teachers to see what kind of feedback was the most beneficial to teachers. The measurement of change in teacher behavior was inferential, that is, it was inferred from a difference between post-interval and pre-interval evaluations. Four conditions were studied,

(a) feedback from students only, (b) from supervisor, i.e., vice-principal, only, (c) from both students and supervisors, and (d) from neither (no feedback). It was found that student feedback led to a positive change among teachers. Supervisor feedback added nothing to this effect when combined with student feedback and when alone, produced change in the direction opposite to the feedback as compared to the no feedback condition. Less experienced teachers showed greater receptivity to student feedback than their more experienced counterparts.  

The last statement is significant in terms of the present study. Younger teachers seem more open to accept student feedback. Such evidence seems to support the contention that the youngest teachers, student teachers, may on the average keep the channels of communication between themselves and their students open. Assuming that the majority of student teachers want to succeed in their first experience, we may hypothesize that they would be even more receptive of pupil feedback than regular teachers. A study comparing teacher behavior change on similar criteria between regular and student teachers appears to be necessary to

---

confirm or deny the above hypothesis.

Now we will leave discussion of pupil evaluations with regular teachers and see what types of studies have been done with pupil evaluations of student teachers.

**Studies of Pupil Evaluation of Student Teachers**

It is difficult to assess the significance of a study if little data are presented. An analytical reader may often be able to determine the causes for weaknesses in a study by studying the data presented. But in a study conducted by Porter one cannot easily determine the causes for failure to gain meaningful evidence because he did not present adequate statistical data with the discussion of his study. Student teachers were evaluated by their pupils after several days when they completed a unit of work. Porter did not tell the reader how many days of teaching preceded the evaluations nor how many days the teaching of the units encompassed. It may have been at the beginning, middle or end of the student teaching experience. There was no follow-up pupil evaluation. Porter was interested in determining how closely do average pupil evaluations agree with those of the cooperating teacher.40

In one table presented, one can see that average pupil evaluations correlated highly with the grade assigned to the unit

---

by the supervising teacher. This study may have been more valuable if each student teacher was evaluated by his class after each of several units. It may have been found that subsequent evaluations would show improvement, possibly because the student teachers received feedback on previous ones.

Veldman and Peck developed a thirty-eight item Pupil Observation Inventory for their study with 554 student teachers. These items were compared with Ryans' presented in his book. The form was administered by the cooperating teachers after the student teachers had completed their required teaching sequences. The purpose of this study was not to provide feedback to the student teachers but to determine if high school pupils, grades 7-12, can reliably evaluate student teachers. The data collected seemed to confirm the authors' hypothesis "that pupils can provide at least as much information regarding teacher characteristics as can expert adult judges on the basis of one or two hours of observation." \(^{41}\)

Remmers and his associates have carried on extensive research at Purdue University in pupil evaluation of teachers. Much of the research relates to the development and use of the Purdue Rating Scale for Instructors. The authors of the scale, \(^{42}\)


\(^{42}\)Veldman and Peck, opcit., p. 354.
G.C. Brandenburg and H.H. Remmers, explain that it was developed "to measure in an objective way the student opinion of the ability of an instructor. All the traits measured are ones which an instructor may with effort alter. Therefore the scale presents an instrument which may be used for the improvement of teaching."

In a study with fifty-seven high school student teachers, Remmers had the pupils evaluate their student teachers on only three items from the Purdue Scale, (1) presentation of the subject matter, (2) stimulating intellectual curiosity, and (3) interest in the subject matter. The evaluations were made after the student teachers had completed their experience of an average eleven week duration. A comparison was made with college students' evaluations of their instructors on the same traits after at least a period of twelve weeks. Remmers found, "It was apparent from the data presented that reliable judgments for single traits may be obtained from high school pupils concerning practice teachers at the end of from 10-12 weeks of acquaintance in a classroom situation."

Another interesting conclusion was found in Remmer's study of halo effect. "It is probable that high school pupils will invest the practice teacher with less halo than college students will.

---


their instructors. This conclusion seems to say that high school pupils are more honest and conscientious in their evaluations of at least their student teachers.

If this study had pupils evaluate their student teachers at different times rather than just at the end, it may have also affected reliability one way or another.

In a study reported by Bowman, the entire Purdue Scale was used. Like Remmers, Bowman obtained pupil evaluations after the terms of student teaching ended. A number of critic teachers were able to observe the student teachers during the term and consequently evaluated them with the pupils. The means of the ratings of two to four critic teachers were correlated with ten to forty-two pupils on each of thirty student teachers. The correlations found were not significantly high. They ranged from \(-.50\) to \(+.47\), averaging at \(-.05\). Such results are inconclusive. They seem to indicate that outside observers do not agree highly with the pupils' views of the student teachers on the traits evaluated.

Having concluded a review of the literature, the author is ready to offer a summary of the major findings on pupil evaluations of teachers.

\(^{45}\) Ibid., p. 630.

\(^{46}\) E.E. Bowman, "Pupil Ratings of Student Teachers," Educational Administration and Supervision, XX(February, 1934), p. 144.
Summary of Research on Pupil Evaluation of Teachers

1. Pupils can reliably evaluate teachers on a variety of forms. The more criteria evaluated, the truer will be the picture of assessment.

2. Pupil evaluations seem to have no unfavorable influence on teacher or pupils if properly administered.

3. Pupil evaluations are most beneficial if they are viewed as personal communication between the pupils and the teacher.

4. Evidence points to the conclusion that the majority of students are honest in their evaluations and comments if they are assured anonymity and no academic reprisal.

5. Pupils are constantly evaluating teachers in discussion with their peers, thus teachers may profit from written reactions.

6. There is some evidence to indicate that teachers can improve in the eyes of their pupils after being evaluated by them and having analyzed their opinions.

7. Student evaluations of teachers should be looked upon as only a single factor in an over-all program of evaluation.

8. There needs to be more replication and comparison of methods of pupil evaluations so as to develop the most reliable and valid ones.

9. So long as the teacher is willing to give pupils an opportunity to comment on anything he likes or dislikes about the
class or teacher, he will get considerable information on any item that is an issue with the pupils regardless of the scaled questions which are included or omitted from the instrument.
CHAPTER III

DESIGN OF THE STUDY

This chapter will include the description of the population selected for the study, the manner in which the pupil feedback form was developed, and the statistical procedures used in the analysis of the data.

Population Sample

The population used in this research consisted of two groups: student teachers and their pupils. All the pupils attended junior or senior high schools within Franklin County, Ohio. Pupil ages ranged from twelve to nineteen years. Over twenty-five hundred pupil evaluation forms were filled out for this study. Over nine hundred pupils filled out both Forms I and II, while over six hundred only filled out Form II.

Student teachers who participated in the study represented the disciplines of Speech, English, History and Science. Twenty-two student teachers were Speech majors, five were Science majors, two were History majors and thirteen were English majors. These student teachers taught one or two classes for a period of six to nine weeks. Nine of the Speech student teachers and twelve of the student teachers who taught English conducted their
field experiences during the spring term, while the remaining twenty-two student teachers taught during the winter term at The Ohio State University.

Since one of the important concerns of this research was to determine the value and affect of early pupil feedback to student teachers, it was necessary to establish both experimental and control groups of pupils. The experimental group of pupils evaluated their student teachers after approximately ten days of teaching and then during the final week of the teaching experience. The control groups consisted of pupils in a second class taught by the same student teachers but who evaluated their student teachers only during the final week.

It was recognized that this control group may not be completely free from the influence of the experimental variable, the early pupil evaluation. A given student teacher may have been made aware of certain communication behaviors practiced in his experimental class that also occurred in the control class. If the behaviors were ones that needed strengthening, then a student teach may have consciously striven for improvement in both classes.

On the other hand, no two classes are exactly alike in groupings of interest in the subject, intellectual abilities, cooperative abilities, age differentiations, or social maturity. A teacher must be aware of the differences in group structures and
individual personalities for each of his classes and make adjustments in his manner of teaching which best fit a particular class.

After student teachers in the experimental group had taught approximately ten days, they asked their pupils to evaluate their effectiveness as communicators. Since the evaluation was viewed in terms of feedback, it was important that the student teachers administer the evaluation themselves. If someone else, such as the cooperating teacher or college supervisor, were to administer the forms, the necessary intimacy in the communication cycle would have been lost. In addition, making the evaluation appear to be a personal correspondence between the pupils and their student teacher eliminated some of the connotation that it was only a grading of the teacher's performance.

The student teachers were able to perform their own analysis of the pupil feedback before sharing the results with anyone else. They were asked to record and average the scale values for each of the first twenty-six items on the forms filled out by the pupils. It was hypothesized that this private analysis would be beneficial and cause some dissonance on behaviors conceived differently by the student teacher and his pupils, as well as reinforce the behaviors that they agreed were effective. Student teachers were not forced to share immediately the information gained from the pupil feedback. If they desired to discuss the results with their cooperating teacher or college supervisor, they were free to do so. The reason
for such a procedure was to eliminate the feeling among student teachers that pupils were grading them or that pupil evaluations would influence the grades obtained in student teaching. In most cases familiar to the author, student teachers were quite willing to discuss the initial pupil evaluations with both cooperating teachers and their college supervisors.

The same procedures were followed in the administration of the second pupil evaluation at the end of the terms. Student teacher were able to see if the pupil perceptions of their communication behaviors changed from the time of the initial evaluations.

The pupil evaluations are designated as Form I and Form II for their sequential order of administration. Form III is the data sheet upon which student teachers recorded a class of pupil evaluations. Form IV is the self-evaluation form filled out by a number of the student teachers. All these forms appear in the Appendix.

In order for the reader to understand the value of the particular form used, an elaboration of its development will be presented.

**Development of the Pupil Evaluation Form**

Preceding the development of a form for pupil evaluation, the author perused the many forms already developed for teacher evaluation. From this study, certain criteria were established
on the nature of the form to be developed.

1. The form should ask for evaluation of behaviors that can be recognized and interpreted by high school pupils.

2. The form should be short enough to be done quickly by pupils, but also, it should be extensive enough to give a realistic picture of a teacher's communication behaviors.

3. The form should exhibit an intimacy in phraseology to encourage pupils to feel that their opinions are really being sought to help the student teacher do better.

4. The behaviors being evaluated should be alterable.

5. There should be space for free comments.

6. The feedback form should be adaptable to the specific scope and subject matter taught.

7. The behaviors being evaluated should be descriptive rather than predictive or prescriptive.

8. The manner of evaluating each behavior should not appear to be secretive in order for the process of evaluation to appear meaningful to both pupils and student teachers.

Considering these criteria, the author initially was able to eliminate a number of evaluation formats such as forced-choice, semantic differential, paired-comparisons, and man-to-man scales. Evaluation, to be a form of feedback, should be straight-forward and semantically clear. An open-faced scale offers such a characteristic. If there is a continuum upon which each behavior
is to be evaluated, the rater can extend the perogative of choice. Since it was important that these pupil evaluations should not be interpreted as a type of grading, the elements of the continuum needed to be stated in other than qualitative terms such as excellent, average, or weak. Since recurring behaviors are being evaluated, it appeared that a continuum presented in terms of quantity of occurrence seemed a possible substitute. If the behaviors are so-stated, the more often they are observed, the more effective will be a teacher's ability to teach, then a scale consisting of frequency of occurrence could be used.

After the type of continuum was decided upon, it was necessary to determine the optimal length of the continuum. Since pupils are evaluating communication behaviors, it was deemed important to discover what others involved in the measurement of communication behaviors had to say about the length of a continuum. In the field of Speech we find that there are many involved in the process of evaluating acts of communication with some type of open-faced rating form. The conclusion Robinson offers about the length of a continuum in evaluating communication behaviors is representative of many in the field who have dealt with the problem of constructing evaluation forms. Robinson contends that no more than five to seven degrees of difference should be used, otherwise a false sense of refinement will occur.¹

Considering the above contention, a seven-point continuum from never to always was adopted. The phrasing of the statements of behaviors were adapted so that any item of the continuum could easily be used to describe the frequency of occurrence of the trait being evaluated.

The final form developed consisted of twenty-five specific communication behaviors and one general effectiveness item. No claim is made that this list is exclusive or all-inclusive. Hopefully it is representative of communication behaviors practiced by most teachers. It is comprehensive enough to present an adequate picture of significant communication behaviors of a student teacher. It could not initially be determined if the behaviors were discrete but such information will be gained from an analysis of the intercorrelations of the items.

Since a pre-test and post-test were called for in the experimental design, it was necessary to develop two comparable forms. To reduce the memory factor and yet have similar behaviors evaluated, the verb tense and order of the items were changed between Form I and Form II. Items 1 to 25 on Form I were reversed on Form II, i.e., Item 1 on Form I became Item 25 on Form II, etc. Past, rather than present tense, was used on Form II. The direction of the continuum remained constant for both forms but verb tense of its elements were changed to past tense.

As will be observed by an initial study of Forms I and II
in the Appendix, they appear to be almost identical. Careful scrutiny will cause one to see the differences. The directions on the use of the form were composed so that they could be read orally to the pupils by the student teacher. The phrasing in personal terms was purposely used to hopefully secure some of the intimacy necessary for honest feedback.

As mentioned, Form III is the data sheet on which each student teacher recorded pupil responses. Upon this sheet, the student teacher could list each pupil's responses in a vertical column and each row would thus represent an entire class's responses on a specific behavior. To obtain the average of the pupils' responses for an item, the student teacher just had to sum across a row and divide by the number of vertical columns. These sheets also facilitated transferral of data to IBM cards. From information gained after administration of the present investigation, it is possible to have pupils record their responses on special answer sheets which can permit direct transfer to IBM cards by machine. Such a process would also make it possible to do a thorough statistical analysis of the data for a student teacher and return this information to him in a relatively short time.

All information obtained from the above study was transferred to IBM cards to make it feasible to analyze the data accurately and in the least amount of computation time.
Processes Used to Interpret the Data

Each IBM card contained one pupil's evaluation of his student teacher on either Form I or Form II. Cards were sorted into various groupings for the purpose of testing the null hypotheses. The 630 Computer at The Ohio State University was used for statistical analysis. Three major programs were used to answer questions raised in the investigation. A series of t-tests were needed to determine if the null hypotheses should be rejected. Intercorrelations of behavior items were needed to determine how related were the behaviors being evaluated. And finally, tests of reliability and validity were necessary to determine the usefulness of the forms used.
 CHAPTER IV

PRESENTATION AND STATISTICAL ANALYSIS OF THE DATA

This chapter is an analysis of the data collected during the study. The first section will deal with the pupils' conceptions of what they had learned from their student teachers. The second analysis will include data related to the null hypotheses. A presentation of the pupil evaluation's intercorrelation of items, reliability and validity will be presented in the third section.

Pupils' Conception of their Learning

In order for the pupil evaluations to give consideration to the specific scope and subject matter taught by the student teachers, the second part of the pupil evaluation forms solicited what the pupils thought they had learned under the tutelage of the student teachers. From information obtained from these queries, a student teacher possibly may gain further insight on his ability to teach.

For research purposes, an analysis of pupil responses on the second section of the evaluation form may enlighten educators about what kind of learning pupils perceive themselves to be gaining. Generally, educators hope they are teaching their pupils
how to think, to form values, to develop standards, and to grasp basic concepts and ideas. Bloom, in his taxonomy, proposed that there are seven levels of thought processes that the schools need to develop: memory, translation, interpretation, application, analysis, synthesis and evaluation.\(^1\) Piaget contended that educators must consciously strive to help pupils progress through four major states of thought development: the states of sensory-motor, the pre-operational, concrete operations, and finally, formal operations.\(^2\)

The above goals of education are theoretical, and unless they are translated into student learnings, they hold limited merit. To determine if pupils will respond in any of the above classifications when asked what they have learned well or poorly, a study was made of a stratified random sample of pupils in English and Speech and Drama. Since the evaluation forms were presented at two different times, it was necessary to choose a representative sample from each testing condition. A sample of four hundred pupils was chosen for each pupil evaluation situation.

Certain procedures were followed in choosing the random groupings. Pupils were separated in terms of subject matter.

\(^1\) Bloom, op.cit.

ie., a distinction was made between students taking a required course and one taking an elective course. The required course was English and the elective courses were Speech and Drama. Subgroupings were made according to age and sex. The basic subgroup consisted of twenty-five pupils of similar sex, age and subject area. Randomness was achieved within these subgroups by selecting from the entire population of forms received according to subject area. Age grouping was done according to chronological age and ranged from fourteen to seventeen years. An example of a subgroup would be twenty-five girls, sixteen years old, enrolled in English classes, who filled out Form I of the pupil evaluation of their student teachers.

The first part of the analysis entailed the counting of the number of responses made by each pupil on part two of the pupil evaluations. Although there was space for only three responses for each of two questions, the pupils were given oral directions that they could add more on the reverse side of the form. Table 1 presents the percentages of 200 pupils per column who gave the indicated number of responses for things they learned or things that needed re-teaching. For each form, the number of pupils in each age grouping giving a specified quantity of responses were first summed. Then this sum was divided by the total number of pupils for each subject grouping, 200, to obtain the listed percentages. In the Appendix are the tables from which
TABLE 1

PERCENTAGES OF PUPILS' FREQUENCY OF RESPONSES ON WHAT WAS OR WAS NOT LEARNED

<table>
<thead>
<tr>
<th>Row</th>
<th>No. of Responses</th>
<th>Column 1</th>
<th>Column 2</th>
<th>Column 3</th>
<th>Column 4</th>
<th>Column 5</th>
<th>Column 6</th>
<th>Column 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>+3</td>
<td>42.5a</td>
<td>67.0</td>
<td>54.8</td>
<td>55.0</td>
<td>66.0</td>
<td>60.5</td>
<td>57.6</td>
</tr>
<tr>
<td>2</td>
<td>+2</td>
<td>17.0</td>
<td>15.0</td>
<td>16.0</td>
<td>15.5</td>
<td>14.0</td>
<td>14.8</td>
<td>15.4</td>
</tr>
<tr>
<td>3</td>
<td>+1</td>
<td>17.5</td>
<td>5.0</td>
<td>9.8</td>
<td>10.5</td>
<td>9.0</td>
<td>9.8</td>
<td>9.8</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>26.0</td>
<td>13.0</td>
<td>19.5</td>
<td>19.0</td>
<td>11.0</td>
<td>15.0</td>
<td>17.5</td>
</tr>
<tr>
<td>5</td>
<td>-3</td>
<td>1.0</td>
<td>4.5</td>
<td>2.8</td>
<td>2.5</td>
<td>3.5</td>
<td>3.0</td>
<td>2.9</td>
</tr>
<tr>
<td>6</td>
<td>-2</td>
<td>5.5</td>
<td>8.5</td>
<td>7.0</td>
<td>2.5</td>
<td>9.0</td>
<td>5.8</td>
<td>6.4</td>
</tr>
<tr>
<td>7</td>
<td>-1</td>
<td>9.5</td>
<td>22.0</td>
<td>15.8</td>
<td>13.5</td>
<td>20.0</td>
<td>16.8</td>
<td>16.3</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>84.0</td>
<td>65.0</td>
<td>74.5</td>
<td>81.5</td>
<td>67.5</td>
<td>74.5</td>
<td>77.0</td>
</tr>
</tbody>
</table>

aComputed with the use of a slide rule.

bThe absolute value indicates the quantity of responses whereas the signs only signify direction of learning. (+) Denotes that pupils listed descriptions of learnings developed and (-) denotes descriptions of subject matter not learned.
Table 1 was calculated.

Each column in Table 1 represents a different subject matter. The Roman numeral preceding each listing identifies the form number. Row 1 presents the percentages of pupils giving three responses, Row 2, the number presenting two responses, etc. Starting with Row 5 is the listing of the percentages of pupils who wrote out things they had not learned very well. These continue through Row 8.

Column 3 presents the average of Columns 1 and 2. In Column 6 will be found the average of Columns 4 and 5. The mean percentages for the four conditions are found in Column 7.

On both Forms I and II, over 50% of the pupils listed two or three things that they had learned well from the student teacher. No significantly greater percentage of responses was found on either form. It was hypothesized that pupils would not be able to list as many learnings on the initial evaluations because of their short exposures of approximately ten days. But the data does not support this hypothesis. Pupils were able to list as many learnings on the first evaluations as on the second ones. Possibly an explanation for these results was that the format restricted further listings. Since space and enumeration allowed for only three listings on both forms, many pupils may have thought that was the maximum they needed to list.

There were no significant differences in percentages between
students required to take the course and those who elected it. Seemingly, just because a pupil is required to take a course, he is not restricted in his listing of learnings that he perceived.

Sex or age dimensions did not seemingly affect the number of responses given by pupils. As seen in the tables in the Appendix, variations were random and not very great.

For all conditions, no greater than one-fourth of the pupils failed to list anything that they had learned. Such a finding seems to indicate the great majority of the pupils in the study were quite willing to inform their student teacher about what they had learned under his tutelage.

The converse of the above finding seemed to be true for the section in which pupils were asked to list learnings that needed re-emphasis or re-teaching. The great majority of the pupils, from two-thirds to four-fifths of them, listed no responses in this section. No greater than 20% were even able to list one learning that needed re-development. The data seems to indicate that pupils were reluctant to list learnings not fully gained or were not conscious of what they had failed to learn. On the other hand, just because a pupil can list something that he has learned does not guarantee that he has. Nevertheless, this analysis seems to point to the conclusion that it is not very profitable to ask students to tell a teacher what they have not learned.
Possibly a teacher can gain from the listing of positive learnings. Areas that he has taught that receive little or no mention may be ones that need further attention.

Classification of Learnings

In the analysis of the perceptions of what pupils thought they had learned, an attempt was also made to classify their responses. Eight categories arose from the analysis. Three definable major groupings seemed evident: skills learned, content learned, and value judgments learned. Each of these categories was divided into specific or general designations. The last two categories were least meaningful. They did not comprise more than 25% of any group's responses. One category listed class discipline behaviors such as learning to keep one's mouth shut while the teacher or another pupil is talking. The eighth category consisted of responses of general subject headings such as poetry, grammar, or oral interpretation.

To calculate the percentages in Table 2, the total possible number of responses was determined for each of the four groupings in Table 1. Each of the columns in Table 2 are similar to those in Table 1. The rows list the percentages of pupils who recorded each type of learning.

Analysis of Table 2 indicates over 60% of the learnings listed were classified as either skills or content learnings. Skills consisted of learning about or doing processes. Content
<table>
<thead>
<tr>
<th>Response Descriptions</th>
<th>I-English</th>
<th>I-Speech</th>
<th>$\bar{X}+2$</th>
<th>II-English</th>
<th>II-Speech</th>
<th>$\bar{X}+5$</th>
<th>$\bar{X},1,2,4,5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specific Skills</td>
<td>6.5&lt;sup&gt;a&lt;/sup&gt;</td>
<td>19.7</td>
<td>13.1</td>
<td>6.5</td>
<td>16.5</td>
<td>11.5</td>
<td>12.3</td>
</tr>
<tr>
<td>General Skills</td>
<td>21.0</td>
<td>22.3</td>
<td>21.6</td>
<td>17.5</td>
<td>23.6</td>
<td>20.5</td>
<td>21.1</td>
</tr>
<tr>
<td>Specific Content</td>
<td>21.8</td>
<td>17.8</td>
<td>19.8</td>
<td>37.0</td>
<td>17.5</td>
<td>26.3</td>
<td>23.5</td>
</tr>
<tr>
<td>General Content</td>
<td>11.4</td>
<td>6.1</td>
<td>8.9</td>
<td>8.7</td>
<td>7.9</td>
<td>8.3</td>
<td>8.6</td>
</tr>
<tr>
<td>Specific Value Judgments</td>
<td>7.7</td>
<td>7.1</td>
<td>7.4</td>
<td>6.5</td>
<td>7.2</td>
<td>6.9</td>
<td>7.1</td>
</tr>
<tr>
<td>General Value Judgments</td>
<td>10.8</td>
<td>2.9</td>
<td>6.9</td>
<td>3.1</td>
<td>.2</td>
<td>1.8</td>
<td>4.3</td>
</tr>
<tr>
<td>Class Discipline</td>
<td>3.4</td>
<td>1.0</td>
<td>2.2</td>
<td>1.9</td>
<td>.6</td>
<td>1.3</td>
<td>1.8</td>
</tr>
<tr>
<td>General Readings</td>
<td>15.9</td>
<td>16.8</td>
<td>16.0</td>
<td>16.0</td>
<td>24.2</td>
<td>20.1</td>
<td>18.2</td>
</tr>
</tbody>
</table>

<sup>a</sup>Computed with the use of a slide rule.
learning consisted of learning about facts, concepts, "truths," etc.. From six to eleven percent of the learning was perceived by the pupils to be the development of values or appreciations, such as, literature is important because it is a mirror of life through the ages.

One group only, Speech and Drama pupils filling out Form II, exceeded twenty percent in listing uninterpretable responses. The data seem to indicate that the majority of pupils can be explicit in listing what they have learned. The explicitness, though, seems confined to the three general categories of skills, content, and values.

Statistical analysis of the second part of the pupil evaluation form was done without the aid of a computer because of the nature of the data. But to analyze the first half of the form in terms of null hypotheses without the aid of a computer would be uneconomical. The computation of t-tests, which were needed to answer many of the null hypotheses, was simplified by procuring a computer program for t-tests. Even a series of forty t-tests can be done in a matter of minutes on a 630 IBM Computer. In the following section will be found data related to each of the null hypotheses.

Each of the null hypotheses will deal with the items in the order presented on Form I. To facilitate reference to specific items, the list of items will be presented below.
1. Studying for this subject has been fun because it is made interesting and exciting.

2. The subject matter has been shown to be of value because much of it can be used in daily life.

3. Questions are asked in a vocabulary that can be understood by all the students.

4. Phrases, such as OK or all right, are not necessarily often repeated.

5. Understanding of student questions or answers is evident because they are related to the on-going discussion.

6. Additional individual help is pleasantly given on any problem not immediately understandable.

7. Certain gestures bring attention to themselves and thus are distracting.

8. Personal appearance and clothing worn are suitable and neat.

9. Ideas are discussed with words that can be understood by everyone.

10. The class is exciting and looked forward to by everyone.

11. Subject matter is easily understood because its learning is not rushed and it is gone into in depth.

12. Homework assignments have been made challenging and interesting and not just busy work.

13. The class is well-behaved and only limited control is necessary to keep it running smoothly.
14. A sense of humor brightens lectures and discussions.
15. It is evident that a lot is known about the topics taught.
16. Homework and in-class assignments are made quite clear.
17. Class discussions are lively, enlightening, and interesting and everyone participates.
18. Tests and grading are fair and do not favor anyone or anything.
19. The amount of work required is reasonable and necessary to get the most out of this class.
20. Student ideas are asked for and given due consideration.
21. In this class, many important ideas are being learned.
22. All that goes on or happens in this class is immediately recognized and effectively dealt with.
23. A pleasing personality is evident.
24. The classroom atmosphere is friendly.
25. A voice is used that is enjoyable to listen to.
26. Considering everything, the teacher does a good job of teaching this class.

Statistical Data on the Null Hypotheses

H1: There is no significant difference in pupil evalua-
tions of student teachers between a first and second evaluation.

Forty student teachers had their pupils evaluate them with the initial and final evaluation forms. This group consisted of seventeen in Speech and Drama, eighteen in English, and five in Science. They taught either during the winter or spring term of 1969 at The Ohio State University. Over 900 pupils evaluated these student teachers each time.

In order to test this hypothesis, it was necessary to determine if pupils, on the average, changed their perceptions of a student teacher from one evaluation to another. The statistical test which will present to the researcher any possible significance between averages is called a t-test for means. Significance is viewed in terms of how many chances out of one hundred the difference in means could be attributed to chance. The minimum level usually used in most research is .05. What this may be interpreted to mean is that the difference in means or averages is less than five in one hundred due to chance. Finding such significance does not explain the reason or cause for the difference in means. There is a high probability that some variable or variables are in effect.

Viewing Hypothesis I in terms of part of the rationale developed for this research, a significant change in the positive direction on a number of communication behaviors may indicate a student teacher made effort to align his perception with those of the students. From analysis of the data, it may be
determined which items most often showed significant change and in what direction. A two-tailed t-test will indicate the direction of significance, that is, positive or negative. Positive significance signifies that the mean of the second sample is significantly greater than the mean of the first sample. Similarly, a negative t-value signifies that the mean of the first sample is significantly greater than the mean of the second sample.

Each of the twenty-six items of the pupil evaluation forms for each student teacher was tested for significance. In order that the same communication behaviors be tested on the two different evaluation forms, the IBM cards were punched with the same items being recorded in the same column, even though the order of the items had been reversed in Form II. In other words, Item 1 on Form I was punched in the same column as Item 25 of Form II. A similar procedure was used in punching all the other reversed ordered items. Only Item 26 on each form remained constant.

Table 3 presents a summary of the number of items each of the forty student teachers showed significant change. Since the change occurred in both direction, separate columns were needed to indicate the number of positive or negative changes. The final column contains the number of items that showed no significant change.

Thirty of the student teachers showed significant changes in communication behaviors as perceived by the pupils on five or more items. Of those thirty, the majority, seventeen, showed signifi-
### TABLE 3
THE TOTAL NUMBER OF CHANGES EACH STUDENT TEACHER MADE AS PERCEIVED BY HIS PUPILS

<table>
<thead>
<tr>
<th>Student Teacher</th>
<th>Positive Changes</th>
<th>Negative Changes</th>
<th>No Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>24</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>2</td>
<td>24</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>3</td>
<td>19</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>9</td>
<td>16</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>16</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>9</td>
<td>16</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>3</td>
<td>23</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>20</td>
<td>6</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>3</td>
<td>21</td>
</tr>
<tr>
<td>17</td>
<td>0</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>18</td>
<td>3</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>19</td>
<td>1</td>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td>20</td>
<td>1</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>21</td>
<td>2</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>22</td>
<td>0</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>23</td>
<td>3</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>24</td>
<td>1</td>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td>25</td>
<td>1</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>26</td>
<td>2</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>27</td>
<td>0</td>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td>28</td>
<td>0</td>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td>29</td>
<td>9</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td>30</td>
<td>10</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>31</td>
<td>0</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>32</td>
<td>3</td>
<td>4</td>
<td>19</td>
</tr>
<tr>
<td>33</td>
<td>2</td>
<td>5</td>
<td>19</td>
</tr>
<tr>
<td>34</td>
<td>1</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>35</td>
<td>6</td>
<td>5</td>
<td>19</td>
</tr>
<tr>
<td>36</td>
<td>2</td>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td>37</td>
<td>5</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>38</td>
<td>1</td>
<td>4</td>
<td>22</td>
</tr>
<tr>
<td>39</td>
<td>0</td>
<td>4</td>
<td>22</td>
</tr>
<tr>
<td>40</td>
<td>3</td>
<td></td>
<td>23</td>
</tr>
</tbody>
</table>

*The minimum level of significance chosen was .05 but many items showed even higher significance of .01 or .001.*
cant changes on ten or more items.

The remaining ten student teachers showed significant change on
from one to five items. Not one of the forty showed less than one
significant change.

Twenty-two of the forty student teachers had more changes in
the negative direction than in the positive direction. For only six
teachers there were found more positive than negative changes.
Only two student teachers had an equal number of significant changes
indicated for both directions, student teacher teachers numbers
19 and 24. All except one of the student teachers who showed more
positive changes were English teachers.

To gain further meaning from the data, an analysis was made
of the number of significant changes that were found for each item
of the evaluation form.

From an analysis of Table 4, one finds that certain items
show significant positive or negative differences with fifteen or
more groups. The items that show this greater number of instances
of significance are Items 2, 4, 7, 8, 9, 19, 22, 23, 24, and 25.
In revising the form for future study, attention will be focused on
reducing the number of items with the reliable items of the above
list serving as a nucleus.

For twenty-four items, a larger number of groups of pupils
evaluated their student teachers significantly lower than on their
initial evaluations. On only five items, numbers 5, 6, 10, 11, and
15, did less than ten groups indicate significant change. Such
TABLE 4
THE NUMBER OF PUPIL GROUPS WHO SHOWED CHANGE ON EACH OF THE EVALUATION ITEMS

<table>
<thead>
<tr>
<th>Item</th>
<th>Positive Change</th>
<th>Negative Change</th>
<th>No Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>6</td>
<td>29</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
<td>6</td>
<td>23</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>10</td>
<td>28</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>3</td>
<td>25</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>6</td>
<td>33</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>6</td>
<td>33</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>17</td>
<td>23</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>15</td>
<td>24</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>9</td>
<td>27</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>7</td>
<td>31</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>3</td>
<td>34</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>7</td>
<td>29</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>9</td>
<td>30</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>6</td>
<td>38</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>13</td>
<td>26</td>
</tr>
<tr>
<td>17</td>
<td>2</td>
<td>11</td>
<td>27</td>
</tr>
<tr>
<td>18</td>
<td>3</td>
<td>11</td>
<td>26</td>
</tr>
<tr>
<td>19</td>
<td>1</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>10</td>
<td>30</td>
</tr>
<tr>
<td>21</td>
<td>1</td>
<td>10</td>
<td>29</td>
</tr>
<tr>
<td>22</td>
<td>1</td>
<td>14</td>
<td>25</td>
</tr>
<tr>
<td>23</td>
<td>1</td>
<td>16</td>
<td>23</td>
</tr>
<tr>
<td>24</td>
<td>1</td>
<td>23</td>
<td>16</td>
</tr>
<tr>
<td>25</td>
<td>1</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>26</td>
<td>1</td>
<td>12</td>
<td>27</td>
</tr>
</tbody>
</table>
findings seem to indicate that pupils become more critical on the second or final evaluation. A halo effect seems to be operating during the initial evaluation. Pupils are either less critical and more protective of their student teachers or are not able to form reliable opinions on an early evaluation. The latter contention can be rejected on the basis of data to be presented later on in this manuscript. It was found that item reliability was significantly high for all but one item on both forms.

From the data presented above, it appears that the null hypothesis can be rejected. Pupils did evaluate their student teachers differently on an early and a final evaluation.

H2: There are no significant differences in final pupil evaluations between student teachers who did and did not take the initial self-evaluations.

A group of 197 pupils was compared with a group composed of 716 pupils. On the basis that self-evaluations may cause the student teacher to become more introspective and cause dissonance, it would seem that there may be more changes occurring between evaluations for student teachers who did an initial self-evaluation. If there are a significant number of changes then, it may be concluded that the use of self-evaluations is important in the framework of evaluation used in this study.

Table 5 presents a summary of the pupil evaluations which did or did not show significant differences. Fifteen of the items showed
### TABLE 5

ITEM SIGNIFICANCE SUMMARY FOR PUPILS OF STUDENT TEACHERS WHO TOOK AN INITIAL SELF-EVALUATION AND THOSE WHO DID NOT

<table>
<thead>
<tr>
<th>Item</th>
<th>Positive Significance</th>
<th>Negative Significance</th>
<th>No Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>-</td>
<td>0.7404</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>0.598</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>0.747</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>-1.6981</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-0.3667</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-1.1563</td>
</tr>
<tr>
<td>7</td>
<td>3.6778^b</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>-1.3289^b</td>
<td>-0.5116</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>-1.3289^b</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>2.3376^d</td>
<td>-</td>
<td>1.4693</td>
</tr>
<tr>
<td>11</td>
<td>-</td>
<td>-</td>
<td>1.4675</td>
</tr>
<tr>
<td>12</td>
<td>-</td>
<td>-</td>
<td>-1.3775</td>
</tr>
<tr>
<td>13</td>
<td>-</td>
<td>-3.2401^c</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>-</td>
<td>-3.777^b</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>-</td>
<td>-</td>
<td>-1.0260</td>
</tr>
<tr>
<td>16</td>
<td>-</td>
<td>-</td>
<td>0.0777</td>
</tr>
<tr>
<td>17</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td>2.6902^d</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td>-</td>
<td>-2.4578^d</td>
<td>-1.0769</td>
</tr>
<tr>
<td>20</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>21</td>
<td>2.5081^d</td>
<td>-</td>
<td>-0.6082</td>
</tr>
<tr>
<td>22</td>
<td>-</td>
<td>-</td>
<td>0.3051</td>
</tr>
<tr>
<td>23</td>
<td>-</td>
<td>-6.1991^b</td>
<td>-</td>
</tr>
<tr>
<td>24</td>
<td>-</td>
<td>-3.3063^c</td>
<td>-</td>
</tr>
<tr>
<td>25</td>
<td>2.6525</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

^a Degrees of Freedom = 911, N1=716, N2=197.

^b Significant at .001, needed: 3.373

^c Significant at .01, needed: 2.590

^d Significant at .05, needed: 1.970
no significant differences. Of the remaining eleven there was almost an even split of significance in the positive and negative directions. If one were to mutually exclude positive and negative significances in pairs, there would be technically only one item of significance left.

The data give too little evidence to reject the null hypothesis. In other words it does not seem to matter very much whether student teachers fill out an initial self-evaluation. If student teachers skip the initial self-evaluations, they will still affect as many changes as ones who filled out an initial self-evaluation.

H3: There is no significant difference between pupil evaluations of student teachers who were evaluated initially and those who were given only a final evaluation.

The groups compared consisted of 913 pupils who did both an initial and final evaluation and only 81 pupils who did a final evaluation. A significantly positive difference on any item would indicate that pupils evaluate student teachers more highly if they do not evaluate them on an early evaluation. On almost half of the items, twelve, pupils evaluated their student teachers significantly higher. On only one item, Item 17, was there significance in the opposite direction. Relating these findings with those found for Hypothesis 1, it seems that the first time students are asked to evaluated their student teachers, either in the beginning or at the end of the experience, they give significantly higher ratings.
Data presented in Table 6 seem to suggest that pupils who do the initial evaluations become more critical of their student teachers than those who do only a final evaluation.

After a re-examination of the key used for the continuums of Forms I and II, it was realized that the key for Form II was not strictly stated in quantitative terms. Instead it was a mixture of qualitative and quantitative designations. At first it was thought that possibly this unintentional switching may have been responsible for the opposite significances found for H1. But further consideration and the data found for H3 suggest a different conclusion. If the difference of the continuums caused a different manner of evaluation by the students, then Table 6 should have shown many more negatively significant items.

With some reservation because of the smaller sample used, the data suggest that the kind of terms used in the continuum does not matter and pupils rate student teachers higher the first time they are given such an opportunity.

Another conclusion the data for H3 and H1 jointly suggest is a new way at looking at pupil evaluations. They may do more for the pupils than they do for the student teachers. Pupil evaluations may serve as a teaching device for the pupils. Pupils may be made aware that effective teaching involves more than a consideration of teacher personality. Once a number of teacher behaviors are identified for the pupils, it seems that they take more cognizance
### Table 6

The summary of item significance of pupils taking only a final evaluation and those taking both initial and final evaluations.

<table>
<thead>
<tr>
<th>Item</th>
<th>Positive Significance</th>
<th>Negative Significance</th>
<th>No Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>-</td>
<td>0.3297</td>
</tr>
<tr>
<td>2</td>
<td>2.7746^b</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2.1201^c</td>
<td>-</td>
<td>0.0111</td>
</tr>
<tr>
<td>4</td>
<td>2.7413^c</td>
<td>-</td>
<td>1.7868</td>
</tr>
<tr>
<td>5</td>
<td>3.1228^c</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>3.0870^c</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>-</td>
<td>0.2054</td>
</tr>
<tr>
<td>10</td>
<td>2.2682^b</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>-</td>
<td>-</td>
<td>0.9635</td>
</tr>
<tr>
<td>12</td>
<td>-</td>
<td>-</td>
<td>1.5361</td>
</tr>
<tr>
<td>13</td>
<td>-</td>
<td>-</td>
<td>0.6375</td>
</tr>
<tr>
<td>14</td>
<td>-</td>
<td>-</td>
<td>1.3288</td>
</tr>
<tr>
<td>15</td>
<td>-</td>
<td>-</td>
<td>0.3731</td>
</tr>
<tr>
<td>16</td>
<td>2.7055^b</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>-</td>
<td>-2.1342</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>2.2533^c</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>4.7158^d</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>2.7024^b</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>-</td>
<td>-</td>
<td>-0.8641</td>
</tr>
<tr>
<td>22</td>
<td>-</td>
<td>-</td>
<td>1.2665</td>
</tr>
<tr>
<td>23</td>
<td>2.1772^c</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>-</td>
<td>-</td>
<td>0.5983</td>
</tr>
<tr>
<td>25</td>
<td>-</td>
<td>-</td>
<td>0.1678</td>
</tr>
<tr>
<td>26</td>
<td>3.9095^d</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

---

^a Degrees of Freedom = 992, N₁ = 913, N₂ = 81

^b Significant at .01, needed: 2.590

^c Significant at .05, needed: 1.970

^d Significant at .001, needed: 3.373
of the occurrence of these behaviors.

Giving two evaluations seem necessary to overcome the halo effect or generally high ratings. On the first evaluations pupils are less critical for possibly two reasons. The process of evaluation may be new to them and therefore the pupils are not actually sure how to react to such an opportunity. Secondly, the pupils may need a first evaluation to point out what behaviors of their teachers they should become more cognizant.

With the large number of significant differences found there appears to be cause to reject the null hypothesis.

H4: There is no significant difference between pupil evaluations of student teachers in required subjects and those in elective subjects.

The subject areas chosen to be compared were pupils in a required course of English and pupils in the elective subjects of Speech and Drama. Since two forms were used, it seemed necessary to test the variable under consideration for each form separately.

Table 7 presents the data of the comparison of groups of pupils in elective subjects with those in a required subject for Form I. Negative significance can be interpreted to mean that pupils in elective subjects evaluated their student teachers significantly higher on an item than did those in the required subject. Positive significance yields the opposite meaning, that being, pupils in the required course evaluated their student teachers significantly higher than those in the elective subjects.
<table>
<thead>
<tr>
<th>Item</th>
<th>Mean N1</th>
<th>Mean N2</th>
<th>t-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.3161</td>
<td>1.4147</td>
<td>-7.9911*</td>
</tr>
<tr>
<td>2</td>
<td>5.0402</td>
<td>1.4170</td>
<td>-4.1207*</td>
</tr>
<tr>
<td>3</td>
<td>5.8103</td>
<td>5.4977</td>
<td>-2.9363**</td>
</tr>
<tr>
<td>4</td>
<td>4.167</td>
<td>1.1728</td>
<td>-1.8447</td>
</tr>
<tr>
<td>5</td>
<td>5.1397</td>
<td>5.2673</td>
<td>-1.5888</td>
</tr>
<tr>
<td>6</td>
<td>5.6437</td>
<td>5.6129</td>
<td>-0.2781</td>
</tr>
<tr>
<td>7</td>
<td>3.1063</td>
<td>3.2442</td>
<td>0.9219</td>
</tr>
<tr>
<td>8</td>
<td>6.2069</td>
<td>6.2212</td>
<td>0.1133</td>
</tr>
<tr>
<td>9</td>
<td>5.9425</td>
<td>5.5138</td>
<td>-1.5054*</td>
</tr>
<tr>
<td>10</td>
<td>5.137</td>
<td>4.977</td>
<td>-1.9811*</td>
</tr>
<tr>
<td>11</td>
<td>5.3619</td>
<td>5.0115</td>
<td>-3.1407**</td>
</tr>
<tr>
<td>12</td>
<td>4.951</td>
<td>4.5968</td>
<td>-2.5899**</td>
</tr>
<tr>
<td>13</td>
<td>5.0172</td>
<td>4.9562</td>
<td>-0.4666</td>
</tr>
<tr>
<td>14</td>
<td>5.7500</td>
<td>5.1055</td>
<td>-2.9574**</td>
</tr>
<tr>
<td>15</td>
<td>5.8534</td>
<td>5.4171</td>
<td>-4.1411*</td>
</tr>
<tr>
<td>16</td>
<td>5.7155</td>
<td>5.2650</td>
<td>-3.3625**</td>
</tr>
<tr>
<td>17</td>
<td>5.3764</td>
<td>4.9816</td>
<td>-3.3625**</td>
</tr>
<tr>
<td>18</td>
<td>5.7114</td>
<td>5.7696</td>
<td>0.2333</td>
</tr>
<tr>
<td>19</td>
<td>5.7845</td>
<td>5.0853</td>
<td>-5.6978*</td>
</tr>
<tr>
<td>20</td>
<td>5.6034</td>
<td>5.7926</td>
<td>1.7997</td>
</tr>
<tr>
<td>21</td>
<td>5.1052</td>
<td>4.9032</td>
<td>-1.6206</td>
</tr>
<tr>
<td>22</td>
<td>5.2213</td>
<td>4.8963</td>
<td>-2.3603**</td>
</tr>
<tr>
<td>23</td>
<td>6.0315</td>
<td>5.8894</td>
<td>-1.4756</td>
</tr>
<tr>
<td>24</td>
<td>5.9626</td>
<td>5.7071</td>
<td>-2.3736***</td>
</tr>
<tr>
<td>25</td>
<td>5.9454</td>
<td>5.5968</td>
<td>-3.1890**</td>
</tr>
<tr>
<td>26</td>
<td>5.8649</td>
<td>5.7559</td>
<td>-1.1088</td>
</tr>
</tbody>
</table>

*Degrees of freedom = 780, N1 = 348 in elective group, N2 = 433 in required group.

*Significant at .001 needed: 3.373

**Significant at .01 needed: 2.617

***Significant at .05 needed: 1.980
On no item was there found any significantly positive difference. But seventeen items out of twenty-six showed high negative significance, most cases at the .01 level or higher. Seemingly the data suggest that student teachers of elective subjects are evaluated more highly than ones in a required subject during an early evaluation. Possibly such a finding can be discussed in terms of student teachers in elective subjects being able to more quickly establish rapport with their pupils.

Attitudes pupils hold toward a course may also affect the way they accept a student teacher. It would seem that students in the required course of English may be expecting more of their student teachers than those in Speech and Drama. Thus English pupils are more critical of their student teacher on an initial evaluation. Even though English pupil ratings were lower than Speech and Drama pupils on the first evaluation, they were still higher than the ratings of all pupils who also filled out a second evaluation.

The differences in pupils evaluating their student teachers in elective and required subjects is significantly reduced by the time of the final evaluation. Table 8 presents the data on the same group of pupils used in Table 7.

Analyzing Table 8, one finds that there are relatively few items on which there are significant differences. For three items pupils in the elective subjects evaluated their student teachers
### TABLE 8

**INDICATION OF SIGNIFICANCE BETWEEN RATINGS IN ELECTIVE AND REQUIRED SUBJECTS FOR FORM II.**

<table>
<thead>
<tr>
<th>Item</th>
<th>Mean $N_1$</th>
<th>Mean $N_2$</th>
<th>t-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.0588</td>
<td>4.9275</td>
<td>-0.8386</td>
</tr>
<tr>
<td>2</td>
<td>5.1294</td>
<td>5.0401</td>
<td>0.0768</td>
</tr>
<tr>
<td>3</td>
<td>5.2510</td>
<td>5.2995</td>
<td>0.3783</td>
</tr>
<tr>
<td>4</td>
<td>5.6131</td>
<td>5.8696</td>
<td>1.4874</td>
</tr>
<tr>
<td>5</td>
<td>5.0311</td>
<td>5.2271</td>
<td>1.5261</td>
</tr>
<tr>
<td>6</td>
<td>5.1961</td>
<td>5.3140</td>
<td>0.8919</td>
</tr>
<tr>
<td>7</td>
<td>5.8275</td>
<td>5.0773</td>
<td>1.2907</td>
</tr>
<tr>
<td>8</td>
<td>5.7725</td>
<td>5.5652</td>
<td>-1.3922</td>
</tr>
<tr>
<td>9</td>
<td>5.3294</td>
<td>5.9611</td>
<td>2.7550</td>
</tr>
<tr>
<td>10</td>
<td>4.6510</td>
<td>4.5797</td>
<td>-0.3958</td>
</tr>
<tr>
<td>11</td>
<td>4.8392</td>
<td>5.1301</td>
<td>2.0783*</td>
</tr>
<tr>
<td>12</td>
<td>4.6627</td>
<td>4.7343</td>
<td>0.4496</td>
</tr>
<tr>
<td>13</td>
<td>4.7137</td>
<td>4.2029</td>
<td>-3.2020**</td>
</tr>
<tr>
<td>14</td>
<td>5.3059</td>
<td>5.2657</td>
<td>-0.2661</td>
</tr>
<tr>
<td>15</td>
<td>5.4314</td>
<td>5.1353</td>
<td>-2.2668*</td>
</tr>
<tr>
<td>16</td>
<td>4.8131</td>
<td>5.0435</td>
<td>1.2564</td>
</tr>
<tr>
<td>17</td>
<td>4.6171</td>
<td>4.7246</td>
<td>0.5174</td>
</tr>
<tr>
<td>18</td>
<td>5.1020</td>
<td>5.5362</td>
<td>2.7068**</td>
</tr>
<tr>
<td>19</td>
<td>4.6157</td>
<td>5.5924</td>
<td>-0.1300</td>
</tr>
<tr>
<td>20</td>
<td>4.9961</td>
<td>5.3575</td>
<td>2.6459**</td>
</tr>
<tr>
<td>21</td>
<td>4.6863</td>
<td>5.6860</td>
<td>-0.0021</td>
</tr>
<tr>
<td>22</td>
<td>4.5725</td>
<td>4.3333</td>
<td>-1.6623</td>
</tr>
<tr>
<td>23</td>
<td>5.3137</td>
<td>5.3816</td>
<td>0.924</td>
</tr>
<tr>
<td>24</td>
<td>4.9765</td>
<td>4.9611</td>
<td>-0.1038</td>
</tr>
<tr>
<td>25</td>
<td>5.0039</td>
<td>5.1014</td>
<td>0.7688</td>
</tr>
<tr>
<td>26</td>
<td>5.3112</td>
<td>5.6522</td>
<td>2.3949*</td>
</tr>
</tbody>
</table>

*Degree of freedom = 460, $N_1 = 255$ in elective group, $N_2 = 207$ in required group.

*Needed for $0.05 = 1.980$

**Needed for $0.01: 2.617$
significantly higher, while on four items pupils in the required subject of English evaluated their student teachers significantly higher.

The factors which caused a larger number of significant differences for the initial evaluations seemed to have almost diminished by the time of the final evaluation. Differences found are few and possibly mutually exclusive. In other words, any significance found has about an equal chance of being negative or positive.

For the initial pupil evaluation of student teachers in required versus elective subjects, there seem to be adequate evidence to reject the null hypothesis. This conclusion is not true when the variable is tested with the second evaluation form. Very few significant differences are found when controlling for elective or required courses and using Form II, the second evaluation.

H5: There is no significant difference between initial self-evaluations and the initial pupil evaluations of communication behaviors.

Studying Table 9, one will find that for most items there were no significant differences. The few items that showed significance did not demonstrate a trend in either direction.

One exception to these findings was Item 26, the general effectiveness statement. It showed a very high significant difference, .001 level.
<table>
<thead>
<tr>
<th>Item</th>
<th>Mean N₁</th>
<th>Mean N₂</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.8126</td>
<td>4.2812</td>
<td>-1.9569</td>
</tr>
<tr>
<td>2</td>
<td>4.6518</td>
<td>4.3750</td>
<td>-0.8320</td>
</tr>
<tr>
<td>3</td>
<td>5.7017</td>
<td>5.4375</td>
<td>-1.0251</td>
</tr>
<tr>
<td>4</td>
<td>4.2786</td>
<td>4.0625</td>
<td>-0.6687</td>
</tr>
<tr>
<td>5</td>
<td>5.3872</td>
<td>5.5625</td>
<td>0.6988</td>
</tr>
<tr>
<td>6</td>
<td>5.6713</td>
<td>5.8125</td>
<td>0.5197</td>
</tr>
<tr>
<td>7</td>
<td>3.0822</td>
<td>2.8750</td>
<td>-0.5683</td>
</tr>
<tr>
<td>8</td>
<td>6.2953</td>
<td>6.1875</td>
<td>-0.4617</td>
</tr>
<tr>
<td>9</td>
<td>5.7604</td>
<td>5.4062</td>
<td>-1.4049</td>
</tr>
<tr>
<td>10</td>
<td>4.7521</td>
<td>4.3125</td>
<td>-1.3575</td>
</tr>
<tr>
<td>11</td>
<td>5.2017</td>
<td>4.1687</td>
<td>-2.6297</td>
</tr>
<tr>
<td>12</td>
<td>4.7674</td>
<td>5.7500</td>
<td>2.9010</td>
</tr>
<tr>
<td>13</td>
<td>5.0056</td>
<td>4.8125</td>
<td>0.5906</td>
</tr>
<tr>
<td>14</td>
<td>5.6374</td>
<td>5.5625</td>
<td>0.2549</td>
</tr>
<tr>
<td>15</td>
<td>5.6176</td>
<td>5.1250</td>
<td>-2.0569***</td>
</tr>
<tr>
<td>16</td>
<td>5.7637</td>
<td>5.3137</td>
<td>-0.4517</td>
</tr>
<tr>
<td>17</td>
<td>5.1755</td>
<td>5.1875</td>
<td>0.0412</td>
</tr>
<tr>
<td>18</td>
<td>5.8381</td>
<td>6.1250</td>
<td>1.0072</td>
</tr>
<tr>
<td>19</td>
<td>5.1429</td>
<td>6.1562</td>
<td>2.3501***</td>
</tr>
<tr>
<td>20</td>
<td>5.7396</td>
<td>5.4375</td>
<td>-1.2023</td>
</tr>
<tr>
<td>21</td>
<td>5.1711</td>
<td>5.7187</td>
<td>-1.7364</td>
</tr>
<tr>
<td>22</td>
<td>5.0432</td>
<td>5.0312</td>
<td>-0.0119</td>
</tr>
<tr>
<td>23</td>
<td>6.0474</td>
<td>5.1687</td>
<td>-2.5301</td>
</tr>
<tr>
<td>24</td>
<td>5.8969</td>
<td>5.9637</td>
<td>0.2859</td>
</tr>
<tr>
<td>25</td>
<td>5.8092</td>
<td>5.5000</td>
<td>-1.1809</td>
</tr>
<tr>
<td>26</td>
<td>5.9150</td>
<td>4.9062</td>
<td>-4.5336***</td>
</tr>
</tbody>
</table>

aDegree of freedom = 713 N₁ = 718, N₂ = 32.

*Significant at .01, needed: 2.610.

**Significant at .05, needed: 1.970.

***Significant at .001, needed: 3.370.
Pupils, designated by $N_1$, rated their student teachers much higher than did the teachers themselves on general effectiveness. Even though student teachers rated themselves relatively high on specific behaviors on Form I, they cast a realistic appraisal or less inflated rating on their over-all ability. Possibly student teachers are, as their pupils, less sophisticated or aware of what actually makes a teacher effective at the beginning of their first experience of teaching.

The data do not offer strong evidence for the rejection of the null hypothesis.

$H_6$: There are no significant differences between student teachers' final self-evaluations and final pupil evaluations.

As done for Hypothesis 5, student teacher self-evaluations were formed into a group. In Table 10, one will find that on only six items is there significance. Of those six, five are positive, that is, the student teachers, Group $N_2$, rated themselves higher than did their pupils, Group $N_1$. Notice should also be given to the fact that only five items had negative t-values, one of those being significant, Item 7. Positive significance can be interpreted to mean that student teachers rated themselves higher than did their pupils.

With the trend toward positive t-values indicated with such a small group, twenty-five student teachers, it may be hypothesized that a larger number of items may show positive significant differences with a larger group of teachers.
<table>
<thead>
<tr>
<th>Item</th>
<th>Mean N₁</th>
<th>Mean N₂</th>
<th>t-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.8173</td>
<td>4.9600</td>
<td>0.4313</td>
</tr>
<tr>
<td>2</td>
<td>4.9966</td>
<td>4.8100</td>
<td>-0.5143</td>
</tr>
<tr>
<td>3</td>
<td>5.2690</td>
<td>5.1800</td>
<td>0.8201</td>
</tr>
<tr>
<td>4</td>
<td>5.7259</td>
<td>5.1200</td>
<td>1.2855</td>
</tr>
<tr>
<td>5</td>
<td>5.0931</td>
<td>5.5600</td>
<td>1.7380</td>
</tr>
<tr>
<td>6</td>
<td>5.3283</td>
<td>6.0100</td>
<td>2.5794*</td>
</tr>
<tr>
<td>7</td>
<td>3.9695</td>
<td>2.8000</td>
<td>-2.8013**</td>
</tr>
<tr>
<td>8</td>
<td>5.5051</td>
<td>6.2800</td>
<td>1.5927</td>
</tr>
<tr>
<td>9</td>
<td>5.1777</td>
<td>5.5600</td>
<td>1.3527</td>
</tr>
<tr>
<td>10</td>
<td>4.5110</td>
<td>4.8000</td>
<td>-0.0843</td>
</tr>
<tr>
<td>11</td>
<td>4.9371</td>
<td>4.4000</td>
<td>-1.7986</td>
</tr>
<tr>
<td>12</td>
<td>1.6071</td>
<td>5.4100</td>
<td>2.5001**</td>
</tr>
<tr>
<td>13</td>
<td>1.7073</td>
<td>4.8800</td>
<td>0.5047</td>
</tr>
<tr>
<td>14</td>
<td>5.3316</td>
<td>3.6000</td>
<td>0.9977</td>
</tr>
<tr>
<td>15</td>
<td>5.3638</td>
<td>5.1000</td>
<td>0.1359</td>
</tr>
<tr>
<td>16</td>
<td>4.8191</td>
<td>5.2100</td>
<td>1.2125</td>
</tr>
<tr>
<td>17</td>
<td>4.6785</td>
<td>5.2100</td>
<td>1.8440</td>
</tr>
<tr>
<td>18</td>
<td>5.2301</td>
<td>6.0800</td>
<td>2.5114*</td>
</tr>
<tr>
<td>19</td>
<td>4.6599</td>
<td>5.8000</td>
<td>3.3220**</td>
</tr>
<tr>
<td>20</td>
<td>5.2103</td>
<td>5.3600</td>
<td>0.1110</td>
</tr>
<tr>
<td>21</td>
<td>4.7090</td>
<td>4.7600</td>
<td>0.1818</td>
</tr>
<tr>
<td>22</td>
<td>4.9975</td>
<td>5.2100</td>
<td>2.41762</td>
</tr>
<tr>
<td>23</td>
<td>5.1027</td>
<td>5.3200</td>
<td>0.2973</td>
</tr>
<tr>
<td>24</td>
<td>4.9476</td>
<td>6.0100</td>
<td>3.4933**</td>
</tr>
<tr>
<td>25</td>
<td>5.0711</td>
<td>5.1100</td>
<td>1.3809</td>
</tr>
<tr>
<td>26</td>
<td>5.3871</td>
<td>5.4800</td>
<td>0.3630</td>
</tr>
</tbody>
</table>

*Degree of freedom = 614, N₁ = 591, N₂ = 25

*Significant at .05, needed: 1.970

**Significant at .01, needed: 2.2610

***Significant at .001, needed: 3.370
Though a trend appears to be developing, there is not enough evidence to reject the null hypothesis.

H7: There are no significant differences between student teachers' initial and final self-evaluations.

According to Table 11, there seems little evidence to reject the null hypothesis. Student teachers did not significantly change their opinions of their communication abilities. They either were not able to perceive any change or felt that their teaching behaviors were only slightly altered.

Instead of the pupils being strongly influenced to maintain their first impressions by an initial evaluation, it was the student teachers who were the ones who seemed bound by their initial evaluations. The data do not suggest why student teachers maintain a stable self-view of their teaching abilities as measured on the forms developed for this research. Possibly student teachers come to the realization that effective teaching is developmental and that student teaching is too short a period to perceive any significant development.

Considering the data found for Hypotheses 5 and 6, it would seem that student teachers do not become more discriminating after both doing an initial self-evaluation and receiving early pupil evaluations. They consistently rate themselves higher than do their pupils.

The data do not warrant the rejection of the null hypothesis.

H8: There are no significant differences between initial
TABLE 11

SUMMARY OF ITEM SIGNIFICANCES BETWEEN INITIAL AND FINAL
STUDENT TEACHER SELF-EVALUATIONS*

<table>
<thead>
<tr>
<th>Item</th>
<th>Mean $N_1$</th>
<th>Mean $N_2$</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.2812</td>
<td>4.9600</td>
<td>2.4462*</td>
</tr>
<tr>
<td>2</td>
<td>4.3750</td>
<td>4.8100</td>
<td>1.5158</td>
</tr>
<tr>
<td>3</td>
<td>5.4375</td>
<td>5.4800</td>
<td>0.1157</td>
</tr>
<tr>
<td>4</td>
<td>5.0625</td>
<td>5.1200</td>
<td>2.9710*</td>
</tr>
<tr>
<td>5</td>
<td>5.5625</td>
<td>5.5600</td>
<td>-0.0082</td>
</tr>
<tr>
<td>6</td>
<td>5.8125</td>
<td>6.0400</td>
<td>0.6622</td>
</tr>
<tr>
<td>7</td>
<td>2.8750**</td>
<td>2.8000</td>
<td>-0.2068</td>
</tr>
<tr>
<td>8</td>
<td>6.1875</td>
<td>6.2800</td>
<td>0.3793</td>
</tr>
<tr>
<td>9</td>
<td>5.4062</td>
<td>5.5600</td>
<td>0.4904</td>
</tr>
<tr>
<td>10</td>
<td>4.3125</td>
<td>1.1800</td>
<td>0.6674</td>
</tr>
<tr>
<td>11</td>
<td>4.1687</td>
<td>4.1000</td>
<td>-0.2014</td>
</tr>
<tr>
<td>12</td>
<td>5.7500</td>
<td>5.4100</td>
<td>-1.0283</td>
</tr>
<tr>
<td>13</td>
<td>4.8125</td>
<td>4.8800</td>
<td>0.1623</td>
</tr>
<tr>
<td>14</td>
<td>5.5625</td>
<td>5.6400</td>
<td>0.2743</td>
</tr>
<tr>
<td>15</td>
<td>5.3437</td>
<td>5.2100</td>
<td>0.9036</td>
</tr>
<tr>
<td>16</td>
<td>5.1875</td>
<td>5.2100</td>
<td>-0.3554</td>
</tr>
<tr>
<td>17</td>
<td>6.1250</td>
<td>6.0800</td>
<td>0.1616</td>
</tr>
<tr>
<td>18</td>
<td>6.1562</td>
<td>5.8000</td>
<td>-0.1623</td>
</tr>
<tr>
<td>19</td>
<td>5.4375</td>
<td>5.3600</td>
<td>-1.5233</td>
</tr>
<tr>
<td>20</td>
<td>4.7187</td>
<td>4.7600</td>
<td>-0.2658</td>
</tr>
<tr>
<td>21</td>
<td>5.0312</td>
<td>5.2100</td>
<td>0.1421</td>
</tr>
<tr>
<td>22</td>
<td>5.1687</td>
<td>5.3200</td>
<td>0.6271</td>
</tr>
<tr>
<td>23</td>
<td>5.9687</td>
<td>6.0100</td>
<td>-0.1759</td>
</tr>
<tr>
<td>24</td>
<td>5.9687</td>
<td>6.0100</td>
<td>0.2696</td>
</tr>
<tr>
<td>25</td>
<td>5.0000</td>
<td>5.1400</td>
<td>-0.2224</td>
</tr>
<tr>
<td>26</td>
<td>4.9062</td>
<td>5.4800</td>
<td>1.9481</td>
</tr>
</tbody>
</table>

*Degree of freedom = 55, $N_1 = 32, N_2 = 25$.

**Significant at .05, needed: 2.0100

*** Bench mark statement where rating is in opposite direction.
pupil evaluations of a student teacher and final evaluations
done by pupils who did not take the initial evaluations.

Referring to Table 12, a negative difference signifies
that the control group evaluated their student teacher higher
than did the experimental group on the initial evaluation.

For all items in Table 12, the control groups evaluated
their student teachers equal to or higher than the experimental
groups. Such evidence also lends more support to the finding
suggested by the data for H3: the continuum descriptions do not
have much bearing on the way pupils rate their student teachers
the first time. No matter what type of continuum is used, pupils
will usually assign the higher numerical values to student teachers
the first time they are given the opportunity to rate them.

Since the differences found were not numerically great,
there is not enough evidence to reject the null hypothesis.

H9: There are no significant differences between final
pupil evaluations of a student teacher and those in the control
class who gave only a final evaluation using Form II.

In Table 13, positive significant differences on an item
can be interpreted to mean that the control groups, which were
the second classes taught by each student teacher, evaluated their
student teachers more highly than the experimental groups. There
does not seem to be any trend in the direction of the items that
were found significantly different for the two groups. This finding
may have some bearing on the question of the type of continuum
<table>
<thead>
<tr>
<th>Item</th>
<th>Positive Difference</th>
<th>Negative Difference</th>
<th>No Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>9</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>2</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>6</td>
<td>16</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>7</td>
<td>18</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>4</td>
<td>21</td>
</tr>
<tr>
<td>13</td>
<td>3</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>6</td>
<td>19</td>
</tr>
<tr>
<td>16</td>
<td>0</td>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td>17</td>
<td>12</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td>9</td>
<td>15</td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>7</td>
<td>17</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>5</td>
<td>21</td>
</tr>
<tr>
<td>21</td>
<td>1</td>
<td>7</td>
<td>18</td>
</tr>
<tr>
<td>22</td>
<td>4</td>
<td>7</td>
<td>19</td>
</tr>
<tr>
<td>23</td>
<td>0</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>24</td>
<td>0</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>25</td>
<td>1</td>
<td>7</td>
<td>18</td>
</tr>
<tr>
<td>26</td>
<td>2</td>
<td>7</td>
<td>17</td>
</tr>
<tr>
<td>Item</td>
<td>Positive Difference</td>
<td>Negative Difference</td>
<td>No Difference</td>
</tr>
<tr>
<td>------</td>
<td>---------------------</td>
<td>---------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1</td>
<td>18</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>3</td>
<td>22</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>3</td>
<td>22</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>3</td>
<td>22</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>5</td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>4</td>
<td>17</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>4</td>
<td>19</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>13</td>
<td>6</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>14</td>
<td>3</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>16</td>
<td>6</td>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>17</td>
<td>2</td>
<td>7</td>
<td>17</td>
</tr>
<tr>
<td>18</td>
<td>3</td>
<td>1</td>
<td>22</td>
</tr>
<tr>
<td>19</td>
<td>6</td>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>2</td>
<td>22</td>
</tr>
<tr>
<td>21</td>
<td>3</td>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>22</td>
<td>1</td>
<td>1</td>
<td>21</td>
</tr>
<tr>
<td>23</td>
<td>3</td>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>24</td>
<td>1</td>
<td>1</td>
<td>18</td>
</tr>
<tr>
<td>25</td>
<td>4</td>
<td>3</td>
<td>19</td>
</tr>
<tr>
<td>26</td>
<td>2</td>
<td>2</td>
<td>22</td>
</tr>
</tbody>
</table>
used by the pupils. Combining this data with that found for
H₃, it seems that it does not matter whether the continuum is
expressed in qualitative or quantitative terms.

The data do not warrant the rejection of the null hypothesis.

H₁₀: There are no significant differences among the
intercorrelations of the items on Form I.

In constructing the evaluation form some attention was
given to categorizing the behaviors used in the list. A larger
number of items was developed so that after administration to
a large group of pupils, statistical analysis would determine
which would be the best ones to keep in a future revision of
the form. Naturally some would be found unreliable and need
to be eliminated.

Intercorrelation analysis would aid in the revision of
the form. From an analysis of the intercorrelation coefficients
found in Table 14, the author was able to determine if the items
were significantly related. The lower the coefficient, the less
related will be an item to any it is compared to.

Originally, eight general categories were developed to include
the twenty-five specific behaviors. The categories with their
corresponding items are as follows:

1. Behaviors dealing with semantics - Items 3, 4, 5, 9, and 16.
2. Behaviors dealing with making the subject matter relevant
to the pupils - Items 2, 12, 20, and 21.
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0</td>
<td>0.58</td>
<td>0.45</td>
<td>0.32</td>
<td>0.21</td>
<td>0.34</td>
<td>0.36</td>
<td>0.23</td>
<td>0.33</td>
<td>0.33</td>
<td>0.43</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0.58</td>
<td>0.45</td>
<td>0.32</td>
<td>0.21</td>
<td>0.34</td>
<td>0.36</td>
<td>0.23</td>
<td>0.33</td>
<td>0.33</td>
<td>0.43</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 1A**

**Intercorrelations of the Items on Fulfil Evaluation Form I**
3. Behaviors dealing with physical attributes or appearance - Items 7, 8, and 25.

4. Behaviors which demonstrate that the teacher is willing to do all that is possible to help the pupil learn - Items 6, 11, and 19.

5. Behaviors which measure a teacher's ability to cope with class discipline - Items 13, 22, and 24.


7. Behaviors that cause the pupils to become excited about learning - Items 1, 10, and 17.

8. Behaviors which show the teachers knowledge of the subject matter - Items 15 and 18.

All items except Items 7 and 19 showed a statistically significant relationship to each other. To be significantly related at the .01 level the coefficients had to be greater than .254. Significantly higher correlations were found for the items within each of the above eight groupings.

Item 7 should have shown negative correlations since pupils were to give ratings in the opposite direction of all the other items. Its extremely low intercorrelation coefficients certify that it was not related to any of the other items but these findings must be related to the yet to be discussed topic of item reliabilities. On both forms Item 7 was found to be very unreliable. Students did
not consistently give their student teachers low values as they should have if they interpreted the item correctly. Since the manner in which item reliabilities and intercorrelations are determined is somewhat similar, the findings seem to support the assertion that Item was an extremely poor and useless item.

The reasons for Item 19 being significantly different from all the other items is not immediately apparent. Possibly the amount of work a pupil is expected to do in a course needs to evolve into another category.

There are differences among the intercorrelations but some further undeveloped statistical test would be needed to determine significant differences. The data do no support rejection of the null hypothesis.

H11: There are no significant differences among the intercorrelations of the items on Form II.

The data obtained from the computer analysis of Form II was almost identical with that obtained for Form I. The inclusion of a table of intercorrelations seemed unnecessarily repetitious of Table 14, so it was left out. Since both forms were identical in the behaviors evaluated, the test of intercorrelations for Form II lend more support for the conclusions developed for H10.

The Pupil Evaluation Forms' Reliabilities

As mentioned in discussing the previous hypothesis, the forms'
reliabilities could affect the interpretations of the data. For instance, if Form I was found to be highly unreliable, one would have to conclude that early pupil evaluation of student teachers is a waste of time. But in Table 15 will be found the reliability coefficients for each item on both forms. Only two items on Form I showed non-significant reliability. If these two items were rephrased or eliminated, one would still have a form with twenty-four significantly reliable statements.

Even though almost all the items on Form I yielded significantly high reliabilities, pupils increased their reliabilities on Form II. Twenty-two items showed higher reliabilities on Form II, while only four showed lower reliabilities. Yet the four items that showed lower reliabilities on Form II were still high enough to maintain the .01 level of significance.

On both forms, Item 7 was shown to be unreliable. The item was poorly worded and actually was an experimental item. Since this was the so-called "bench mark" statement, the data suggest that is is not a good idea to reverse the sense of the continuum for one item. It raises more possibilities of confusion than serving the purpose of indicating face validity to the form.

Although Item 4 did not appear significantly reliable on Form I, it became significantly reliable on Form II. Possibly pupils had not been able to detect a distracting vocal mannerism in the early evaluation, but with longer exposure, they were able
## TABLE 15

THE PUPIL EVALUATION FORMS' RELIABILITIES ON EACH ITEM

<table>
<thead>
<tr>
<th>Item</th>
<th>r-Form I b</th>
<th>r-Form II</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.5105*</td>
<td>.6876*</td>
</tr>
<tr>
<td>2</td>
<td>.3715*</td>
<td>.6507*</td>
</tr>
<tr>
<td>3</td>
<td>.3607*</td>
<td>.6310*</td>
</tr>
<tr>
<td>4</td>
<td>.1059</td>
<td>.6193*</td>
</tr>
<tr>
<td>5</td>
<td>.8087*</td>
<td>.6180*</td>
</tr>
<tr>
<td>6</td>
<td>.5057*</td>
<td>.6213*</td>
</tr>
<tr>
<td>7</td>
<td>.1035</td>
<td>.1911</td>
</tr>
<tr>
<td>8</td>
<td>.5786*</td>
<td>.1730*</td>
</tr>
<tr>
<td>9</td>
<td>.1616*</td>
<td>.5633*</td>
</tr>
<tr>
<td>10</td>
<td>.5261*</td>
<td>.6177*</td>
</tr>
<tr>
<td>11</td>
<td>.6788*</td>
<td>.6133*</td>
</tr>
<tr>
<td>12</td>
<td>.8274*</td>
<td>.6109*</td>
</tr>
<tr>
<td>13</td>
<td>.3942*</td>
<td>.6009*</td>
</tr>
<tr>
<td>14</td>
<td>.3135*</td>
<td>.5070*</td>
</tr>
<tr>
<td>15</td>
<td>.8413*</td>
<td>.5810*</td>
</tr>
<tr>
<td>16</td>
<td>.5102*</td>
<td>.5186*</td>
</tr>
<tr>
<td>17</td>
<td>.7177*</td>
<td>.6010*</td>
</tr>
<tr>
<td>18</td>
<td>.4831*</td>
<td>.5830*</td>
</tr>
<tr>
<td>19</td>
<td>.7124*</td>
<td>.3270*</td>
</tr>
<tr>
<td>20</td>
<td>.5238*</td>
<td>.5370*</td>
</tr>
<tr>
<td>21</td>
<td>.5683*</td>
<td>.5928*</td>
</tr>
<tr>
<td>22</td>
<td>.8064*</td>
<td>.4693*</td>
</tr>
<tr>
<td>23</td>
<td>.8858*</td>
<td>.6716*</td>
</tr>
<tr>
<td>24</td>
<td>.5076*</td>
<td>.4838*</td>
</tr>
<tr>
<td>25</td>
<td>.5466*</td>
<td>.6556*</td>
</tr>
<tr>
<td>26</td>
<td>.6164*</td>
<td>.7110*</td>
</tr>
</tbody>
</table>

*aN₁ for Form I = 916, N₂ for Form II = 759.

*Significant at .01, needed: .25L.

bReliability formula used:

\[
 r = \left( \sigma_t^2 - \sum_{i=1}^{n} \sigma x_i^2 + \sum_{i=1}^{n} r_i \sigma x_i^2 \right) / \sigma t^2 \\
 n = \text{number of items}
\]

\[
 r_t = \text{scale reliability}
\]

\[
 \sigma x_i^2 = \text{variance of item } x_i
\]

\[
 \sigma t^2 = \text{variance of scale total } "T"
\]

\[
 r_i = \text{item reliability}
\]

103
to determine if the student teacher repeated words or phrases unnecessarily.
CHAPTER V

ANALYSIS OF THE RESEARCH’S DATA

This chapter will adhere to discussion of data in the same order found in the previous chapter. Such analysis should facilitate any reference to data of specific tables.

Pupils Conception of Their Learning

The purpose of including the second section on the pupil evaluation forms was to solicit feedback related to the learning occurring under the tutelage of the student teacher. If pupils can perceive what they are learning, then it might follow that what they say they have learned and what they actually have learned would be identical. Since there would be little reason for a pupil to lie on an anonymous questionnaire, we must accept their perception of learning as truthful.

If the responses are specific and analytical, as opposed to being topical and superficial, it would seem that we could put more credence on the responses. In the classification of responses, two categories consisted of the latter type of responses. For three groups, not more than 20% of the responses fill in the less meaningful categories of class discipline or general subject headings. The remaining group wrote about 25% of their responses within these categories. In other words from 75-80% of the responses given by pupils were meaningful.
"Meaningful" denotes that responses were related to what the student teacher was teaching and could be classified into three major categories: skills, content and value judgments.

Other methods of classification were tried, such as using Bloom's taxonomy of thought processes. But many of the seven processes can be related to the triad of classification chosen. Memory, translation, and interpretation are processes that can easily be related to the learning of the content of the course. Application, analysis, and synthesis can be assimilated in the learning about skills or the actual development of skills. Evaluation, the seventh and most complex level, can be directly related to the learning of value judgments.

Not only did pupils list a majority of meaningful responses but they were also able to list more than one. Over 50% of the pupils listed two or three descriptions of learnings. On early evaluations, pupils listed as many responses as they were able to list on Form II four to six weeks later. It will be necessary to do further experimentation to determine if given more space, pupils would list a greater number of learning descriptions after a longer exposure to the student teacher.

One would think that pupils could give many more descriptions of learning on a final evaluation of a student teacher than on an early evaluation. The analysis of the number of responses given on Form I and Form II did not support this supposition.
The dimension of sex did not seem to affect the number or types of responses given by pupils. The mental age differences between boys and girls are leveling out during the years under study, fourteen to seventeen years. Girls of the same chronological age as boys do not exhibit, as they do in pre-adolescence, significantly different mental ages.

With pupils passing from adolescence in the approximate years under study, it would seem feasible to predict they would become more conscious of what they are learning. Analysis of the types of responses made by sixteen and seventeen year olds do not show any significant differences from those made by fourteen and fifteen year olds. Juniors and Seniors in high school are thought to exhibit a maturity not possessed by freshman and sophomores. The data seem to indicate this maturity difference does not exist in an understanding of what pupils think they are learning.

When asked to list what areas needed re-teaching or re-examination, pupils generally failed to give any responses. A typical response to this query was, "If I (the pupil) did not learn something, how could I describe it?" The original intent of the question was to obtain from the pupils their opinions on subject areas that were covered poorly or too quickly. It would seem a re-phrasing of the question or more explicit directions to the pupils is needed if the desired type of responses are to be obtained.

The educational value of the above findings seem related to
a means of determining teaching effectiveness of a specific subject matter. Teachers may be able to determine what pupils are learning from something other than the traditional achievement examinations. From an analysis of his pupils' responses, a teacher, either a regular or a student teacher, may be better able to determine the type of learnings that are having the strongest impact on his pupils. This approach is more direct and much easier to interpret. A teacher must infer from the written examination questions what type of learning is being measured.

If the forms upon which pupil opinions are solicited are anonymous, the opinions given possess at least face validity. Pupils have no reasons, such as grade pressures, to distort their opinions. Anonymity may have been one of the reasons for at least 75% meaningful responses on the pupil evaluation forms.

Periodically teachers and pupils need to assess what has been learned. Normally this is done through examining the pupils with tests. Because of the difficulty most teachers have in developing consistently valid and reliable tests, it may be equally profitable to obtain feedback of learning incurred by an open questionnaire, which is to be anonymous and non-graded. From the present study, it would be suggested that more than three responses be solicited.

Pupil Evaluation as Feedback

The rationale developed for this research was that early pupil evaluation would serve as meaningful feedback to student
teachers. If pupils informed the student teacher that he was less effective on a number of communication traits, then, according to the role of feedback in the communication cycle, the student teacher would take additional steps to correct these weaknesses which caused less than full comprehension of the messages.

The data procured for H1 seem to suggest a different rationale. Pupils, after the first opportunity to evaluate their student teachers become more critical of their behaviors. On the first evaluation they will reliably rate student teachers higher than on a later one. The initial evaluation seemed to have made the pupils become more aware of the communication behaviors occurring in the classroom. But, this awareness was not likewise developed by the student teachers. H7 presented data suggesting that student teachers do not significantly alter their perceptions of competence as a teacher-communicator.

Possibly making pupils more sensitive to the student teacher's communication behaviors is just as important as making the student teacher aware of the opinions of the pupils.

Significance in the opposite direction as hypothesized does not necessarily indicate that the pupil evaluations did not serve as feedback mechanisms. Student teachers may have gained deeper understandings of how their pupils perceived their teaching behaviors. But an additional procedure is needed to determine the value of the pupil evaluations as feedback. Student teachers should be asked for a written analysis of what they obtained from their study of the pupil evaluations. Oral comments of this nature were made to the author by
his own student teachers. As discussed previously, pupil evaluations were to be intimate and personal communications between the student teacher and his pupils. If student teachers were asked to share all their thoughts of the analysis of the pupil evaluations with an outside party such as the researcher, much of the intimate nature of the process would probably be lost.

The data does not suggest general improvement for the student teachers in the study, at least as perceived by the pupils. But it was found that the pupils evaluate a student teacher more highly on the first evaluation no matter when it is given, early or late, H3. Possibly if pupils evaluated a teacher quite early just to overcome the positive halo effect, and then gave a second evaluation in the middle of the student teaching experience, one may have been able to detect some sort of improvement by the final evaluation. The shortness of the ten week student teaching experience may also have lessened the possibility of observed improvement. Because of the shortness of a student teacher's actual teaching time, it was not possible to inject a pupil evaluation into the middle of the experience without making the process seem too repetitive. This would be especially be true if student teachers were evaluated in more than one class.
Pupil Evaluation and Dissonance

A number of hypotheses were developed to test the affects of possible dissonance incurred by self- and pupil evaluations of student teachers. Student teachers did not significantly alter their self-evaluations from the initial to the second one, H7. Since pupils evaluated their student teachers generally high in the initial evaluation, little if any, dissonance was aroused, H1. Particular note can be given to the data for H5 in which was found that there was not a large number of significant differences between initial pupil evaluations and the student teachers first self-evaluations made immediately prior to the pupil evaluations with Form I. On only one item was there significantly high difference, the general effectiveness statement, Item 26. Pupils rated their student teachers much higher than did the student teachers rate themselves. Student teachers may have interpreted these higher general effectiveness ratings by pupils to mean that they were really performing quite satisfactorily as a teacher and that slight disagreements in more than one specific behavior statement were not really a detriment to the teacher's communication abilities.

In studying the data for H2, it was found that student teachers were not evaluated differently on the second evaluation just because they filled out an early self-evaluation. The possible dissonance producing affect of comparing one's self-evaluation with pupils' evaluations of an early evaluation appears difficult to determine.
from the data.

Considering the data which tested the value of self-evaluations, one would have to conclude that their use was of little or no value in the format used for this study.

**Pupil Evaluations and Subject Matter**

Supposedly in education there is a dichotomy between required and elective subjects. Required subjects are those prescribed by teachers and administrators to be essential to the development of the child. Supposedly elective subjects are not essential to a child's development, but they provide a means to study interests not covered in the curricula of required subjects. If a student enrolls in an elective subject, it is often assumed that his attitude toward the subject matter is much different from his attitude to a required subject. Since a pupil has an initially favorable interest in an elective subject, he may be of the volition to evaluate anything about the elective course more highly than he would for a required course.

Within this study were found two large groups for comparison of the required-elective dichotomy. Pupil taking Speech and Drama courses were classified as pupils in elective courses, while English pupils were classified as pupils taking a required course. Since these two groups of pupils may evaluate their student teachers differently on early and final evaluations, separate analyses were
The data from Form I yielded some interesting results. On Form I pupils in the elective subjects evaluated their student teachers significantly higher than those in the required subjects.

These differences may possibly be explained in terms of pupils' interests in the subject matter. Since pupils in the elective subjects generally already have a special interest in the subject matter, they may be less critical of the initial performance of the student teacher. This special interest may not generally be held by the pupils in required courses and they may feel with a change of teachers, will also come a better presentation of the subject matter.

Even though pupils in elective subjects evaluated their student teachers significantly higher than those in required subjects, it was observed that the pupils in the required subjects evaluated their student teachers relatively high, four and above.

The difference found between the initial evaluations of pupils in elective and required subjects seemed to have diminished by the time of the second or final evaluation. As found in the data for Form I, pupils become more critical during a second evaluation of their student teachers. "Critical" denotes that pupils evaluated their student teachers with lower average ratings than they did on the initial evaluations. Any differences possibly attributed to subject matter seem negligible. Pupils appear to
become more discriminating when given a second opportunity to evaluate their student teachers.

The forms developed for this study appear usable by both pupils in required and elective subjects, at least those dealt with in this study, English, Speech and Drama. The data do not suggest any causes for the significant differences found between the initial pupil evaluations for the variable of elective versus required subject matter. A more probing study needs to consider the differences and isolate the possible causes.

In this study, the major concern was whether the forms used could be adapted to more than one type of course. At least Form II, was found to be equally usable in either an elective or required subject. Since both forms were found to be highly reliable, the differences in results with the use of Form I have to be assigned to something other than the format used.

**Pupil Evaluation and Its Format**

The format used for pupil evaluation was essentially an open-faced evaluation of communication behaviors, using a seven-point continuum. The behaviors evaluated were chosen on the basis that they were somewhat discrete, meaningful, and necessary to develop a comprehensive view of a student teacher's image as a communicator. From analysis of each item's reliability on Forms I and II, in which each item was compared with every other item, it seems reasonably safe to propose that most items received consistent pupil responses.
Only one item, Item 7, was found to be unreliable on both forms. With its poor phrasing and opposite direction of interpretation, pupils became confused about its meaning. The only other item which was found to be unreliable on one form, Item 4 on Form I, showed significant reliability on Form II. Phrasing may have been responsible for the unreliableness of the item on Form I because "not" was poorly placed in the statement.

A re-examination of the continuum, caused the author to realize that on Form II, the descriptions were not all quantitative. Instead, there was a mixture of quantitative and qualitative descriptions. Initially it was thought that this changing of the continuum might be responsible for obtaining opposite significances for H1. But data for H3, which compared the first uses of Form I and Form II by groups of pupils, show that ratings were high and few oppositely significant differences were found. With some reservations because of the smaller sample size, the data suggest that the kind of terms used in the continuum do not matter. Using either type of continuum, or mixing qualitative and quantitative terms, the pupils rated their student teachers quite high the first time they were given such an opportunity.

Certain items on the two forms appeared to show significant differences more often than others in the data for H1. Pupils seemingly are more discriminating on those items and thus in any revision of the forms, careful consideration must be given to those items. The
items were Items 2, 4, 7, 8, 9, 19, 22, 23, 24, and 25. Two of these items have to be eliminated because of their statistical unreliability. Item 4 was found unreliable on Form I and Item 7 was found unreliable on both forms.

Analysis of the intercorrelations of the items suggest that most of them are somewhat related, but some more so than others. Only Item 19 in Category 4 did not seem related to behavior statements judged to be similar. The data suggest a much shorter form with from eight to nine behavioral statements to be evaluated. In order to obtain from pupils different aspects of each of these behaviors, it is suggested that upon rating each behavior, the pupil be asked to give reasons for his rating. The pupils might be asked to list some of the teacher's actions which caused them to chose a certain rating.

Before any strong generalizations can be made, more pupils need to use the pupil evaluations. Pupils in other subjects such as Mathematics, Social Studies, or Business Education need to use the forms.

This study was conducted as a preliminary research to help the author to determine weaknesses in both his rationale and methodology. As was seen in the presentation of the statistical data many of the researcher's hypotheses were not supported. Consequently the results have to be qualified in terms of any generalizations made.
The format employed seemed defensible in that pupil evaluations on both forms were highly reliable. Whether this high reliability can be attributed solely to specific items used, cannot be determined at this time. It would seem from a study of free response evaluations obtained in a pilot study that the prepared form was necessary. On the free response evaluations, pupils were generally nebulous or vague, too brief, and often wrote evaluations of the student teachers' personalities. The prepared forms did provide for free responses in addition to the evaluation of specific behaviors. Pupils usually wrote comments which explained the reasons for low or high ratings given to the student teachers. But the majority of the pupils failed to write free comments. The data obtained in this study seem to support the findings of a number of studies discussed in Chapter II, in which a prepared form will obtain reliable evaluations.

Possibly a prepared list of behavioral statements may also make the pupils more aware of these behaviors so that on a second evaluation covering the same traits, they will be able to give a clearer perception of the occurrence of the behaviors. At least that is the conclusion the data seem to suggest.

An Over-All Analysis of Pupil Evaluation

The original assumptions upon which this research was based need further study. Some of the data supported views in opposition
to those held by the author. For this study, pupil evaluations
did not appear to serve as a type of feedback which would cause
improvement of the communication behaviors of student teachers.
Instead, the initial evaluations seemed to have had a greater affect
on the pupils. Since pupils were reliable in filling out early
evaluations, it may be proposed that their higher than average
evaluations actually made them become more aware of their student
teachers' behaviors being evaluated. Thus on a second evaluation,
the pupils became more critical of their student teachers and
probably evaluated them more realistically.

This finding may be important. If pupils are conscious of
what a teacher is doing, they will not be amenable to poor teaching.
Making the pupils critical of teaching may in the long run cause
teachers to improve. What actually needs to be done is to have
a follow-up study on the student teacher who participated in this
study. Since pupils became more critical on the final evaluation,
maybe this feedback would cause the dissonance discussed in the
rationale of the study. The next opportunity of teaching for the
student teachers may show the positive gains in communication abilities
not found in this study. It was found that pupils evaluate student
teachers with higher ratings the first time they are given the
opportunity to evaluate. Then a follow-up study would have to compare
a second pupil evaluation with the second evaluation obtained during
student teaching. This comparison would determine if the student
teaching final evaluation acted as feedback and caused the student teacher to make a conscious effort to improve in areas of weakness, as perceived by the pupils.

If the format of pupil evaluations was less than effective in a number of situations, the process was seemingly important. In the author's own experience, his student teachers seemed to have gained from the process. They seemed more open to criticism given by myself or the cooperating teacher as compared with groups of student teachers who did not use pupil evaluations.

Maybe it is psychologically good that pupils evaluated their student teachers quite high on the first evaluations. Since student teachers are apprehensive about their abilities as teachers in the beginning, high early pupil evaluations may serve as a morale booster or confidence builder. In that way they act as a positive form of feedback. The early evaluations do not pin-point areas of weakness but generally positive or high ratings may be necessary for the student teacher to progress.

An analogue to the above contention can be found in the teaching of public speaking. Initially it is important to get the novice speaker on his feet and comfortable in front of an audience. To gain composure and confidence, the speaker must be given positive encouragement and not be criticized in too many specific behaviors. A teacher must build up the speaker's tolerance to receiving negative criticism in a gradual manner.
The student teacher is similarly getting accustomed to his audience, the pupils. For him to get discouraged to the point of giving up seems possible if he receives too much negative criticism. He needs positive reinforcement from the audience, the pupils. It seems that pupils are more precocious than we give them credit, they seem to give student teachers encouragement on an early evaluation. Once they have given this note of encouragement, the pupils seem to expect the teacher to respond by doing a better job of teaching. Pupils in this study did not rate their student teachers as high or higher on the second evaluations.
CHAPTER VI

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

The purpose of this study was to examine the process of pupil evaluation of student teachers. The primary question studied was, "Will early pupil evaluation of student teachers make more effective communicators of those student teachers, as perceived by the pupils?" With the many conflicting findings in previous research, it was hoped that this study would contribute some meaningful findings for further comprehensive research.

To accomplish this task and to obtain data, a pupil evaluation form was developed. This form consisted of two parts, twenty-six statements to be rated along a seven-point continuum and a section in which pupils were asked to list what they had specifically learned or not learned. Since the process of self-evaluation seemed strongly related in the rationale developed, a separate self-evaluation form was developed for student teachers. Essentially, only the phraseology of statements varied from that of the pupil evaluation forms. The behaviors evaluated on both forms were identical.

Summary of Procedure

The research procedures involved (1) a survey of the literature
to gain background and direction, (2) the development of a form for pupil evaluation and one for student teacher self-evaluation, (3) distributing pupil evaluations to a number of student teachers in four subject areas, (4) processing the returned pupil evaluation forms for transferral to IBM cards through keypunch processing, (5) statistical treatment and analysis of the data, and (6) presenting the data in tabular and graphic form.

The assumptions investigated were as follows:

1. Pupils can reliably evaluate their student teachers as early as ten days after they begin teaching. (Q1)

2. The pupil evaluation forms developed will present an adequate picture of the total teaching situation. (Q2)

3. As a result of early pupil evaluations, student teachers improve, in the abilities initially evaluated low, by the end of their experience. (Q3)

4. A student teacher will change his self-perception of himself as a teacher by the end of the student teaching experience. (Q4)

5. Pupils will be discriminating in their ratings of student teachers. (Q5)

6. Pupils will be able to give meaningful lists of what they are learning. (Q6)

**The Statistical Treatment**

In order to test the assumptions of this study the following
statistical procedures were used:

Sums, means, sums of squares, standard deviations, and frequencies were determined by The Ohio State Computer Center and the use of IBM 630 and 1620 computers. These statistics were used in the computation of the $t$-values to determine statistical significance at the minimum five per cent level of confidence.

A simple item analysis adapted to the IBM 1620 computer was used to determine item reliabilities.

A program was secured for use with the IBM 630 computer to determine item intercorrelations.

**Findings**

A summary of the investigator's major findings from the statistical data is as follows:

1. Pupils could reliably evaluate their student teachers after only ten days of initial teaching. (Q1)
2. Pupils evaluated their student teachers even more reliably at the end of the student teacher's experience. (Q1)
3. Pupils seem to become more critical at the time of a second evaluation and do not generally rate student teachers as high as they did with the first evaluations. (Q3, Q5, H1)
4. Pupil evaluations seem to have no affect on the way student teachers evaluated themselves. The second self-evaluations of student teachers are not significantly different from their first ones. (Q4, H7)
5. Pupils evaluated their student teachers with high ratings the first time they participated in the process. Pupils who did not evaluate their student teachers initially gave them high ratings on Form II at the end of the experience. (Q5, H1, H3)

6. When asked what they had learned, a majority of the pupils were willing and able to list a number of meaningful descriptions. (Q6)

6a. Older pupils in the group under study, seventeen year olds, did not list a greater number of meaningful responses than did the younger pupils, the fourteen year olds. (Q6)

7. Pupils in elective classes seem to evaluate their student teachers higher on specific traits than those in required classes on Form I. These differences were not found on the second evaluations taken by the same two groups. (H4)

7a. The pupil evaluation form appeared adaptable to more than one subject area. Possibly the evaluation of communication abilities facilitate this wide use. (H4)

8. Student teacher initial self-evaluations are high and not significantly different from the high initial pupil evaluations. (H5)

9. Using a quantitative or qualitative continuum may not disrupt the way pupils rate teachers. (H3, H8)

Some findings that were not specifically attributable to statistical data were:

1. Student teachers generally felt that they profited from
early pupil evaluation. This finding was derived from open discussion with over half of the student teachers who partook in the study.

2. The transferral of individual pupil evaluation sheets to the data sheets was very laborious and time-consuming. A special answer sheet could have been used which would have permitted machine transferral of responses to IBM cards.

3. Inserting an item written in the opposite sense or direction of all the other items creates a very unreliable item.

4. A larger control group which do not fill out the initial evaluations was needed. Because of the limited sample of student teachers available during the running of the study, it was decided that a larger experimental group was needed to gain more representative results.

5. Student teachers might have been helped more if basic statistical data for each evaluation was computer processed and returned to them quickly.

In summary, some of the original assumptions were supported, while others, though not refuted, lead the author to a re-interpretation of the affects of pupil evaluation of student teachers. Seemingly, pupils are more affected by early evaluations of student teachers than are student teachers. The measuring of the affects on student teachers was not determined and remains to be discovered in future research.
Conclusions

1. Ten days after a student teacher has begun teaching is not too early to have pupils evaluated them reliably. (Q1)

2. An early evaluation seems necessary to introduce pupils to the process. They seem to become more critical after the first evaluation. (H1)

3. Early self-evaluation by the student teachers do not seem to cause them to change more behaviors. The value, if any, of student teacher self-evaluation are not supported by the data. (H2)

4. The type of continuum used may not have any bearing on the manner in which numerical ratings are assigned to student teachers by their pupils. (H3, H8)

4a. An early pupil evaluation seems to act as a teaching device for pupils. Behaviors, not related to personality characteristics, are brought to the attention of the pupils. Consequently, by the time of the second evaluation, pupils have knowingly become more discriminating about the behaviors listed on the first evaluation. Pupils are able to evaluate their student teachers more realistically on the second evaluation. (H3, H1)

5. On early evaluations, pupils in elective subjects evaluated their student teachers significantly higher than pupils taking the required subject. There seems to be more of an open mind existing among pupils in elective subjects on early evaluations. Pupils in elective subjects appear more adaptable to change of
teaching technique and teachers. This difference in rating student teachers diminishes by the time of the second evaluation. Pupils in required subjects as well as elective subjects become more discriminating and probably more realistic on their second rating of student teachers. (H4)

6. Student teachers appear as unsophisticated as pupils on their early self-evaluations. On specific behaviors, student teachers seem to have an inflated view of their abilities. Although student teachers generally receive average grades in academic courses, they seem to have thought that they were significantly above average as a teacher even after only a short exposure to teaching. (H5)

7. Student teachers maintain a stable high opinion of their teaching abilities, even after a quarter of student teaching. Pupils change their opinions on the behaviors of their student teachers but student teachers do not significantly change their opinions of themselves. Pupils appear to become more discriminating on a second evaluation. In that student teachers do not change their self-perceptions, one may contend that this finding lends credence to the proposition that behavior of individuals does not change quickly. But since both early and late self-evaluations of student teachers were significantly high, the more probable conclusion is that a quarter of student teaching does not appear to make student teachers more realistic about their own conception of their teaching abilities. (H6, H7)
8. It may not be necessary to have a student teacher evaluated in more than one class. There seems to be some carry-over from the early evaluation in the other class. Since there was the possibility of a second class of a student teacher not remaining completely free from the influence of another class's early evaluation, it cannot be concluded that first evaluations are not always high as found in the data for H1 and H3. Instead, one can only conclude that two classes of a student teacher evaluate him similarly with Form II, no matter if they do or do not fill out an early evaluation. (H9)

9. Most items on the evaluation forms seem somewhat related, some more so than others. The number of items may be appreciably be reduced to eight or nine behavior descriptions. In order to obtain different or specific facets of each of these behaviors, pupils should be asked to list reasons why the numerical value was assigned to the given statement. Pupils could list specific teacher actions or practices related to the behavior described. (H10, H11)

10. If asked, pupils will elaborate on what they think they have learned. To relate an evaluation to a specific subject matter, it would seem practical to request pupils to describe in a list what they have learned. The space left for such a section should not be prohibitive or restrictive. Pupils may be encouraged to list a minimum number of learnings, but this minimum should not also be the maximum. In other words, pupils may be asked to list
at least three responses, but be given six or more spaces to record the descriptions.

11. Introducing student teachers to the process of pupil evaluation seems particularly adaptable to student teaching in which they have people to consult familiar with the class who did the evaluation.

12. The study was valuable because it yielded results which not only supported some of the original assumptions but also suggested a need for re-consideration of the others. Weaknesses in design became apparent as the study progressed. Much data was collected and analyzed so that the investigator learned the methodology of programming data for a computer. As mentioned in the above text, new approaches have been developed as a result of the analysis of the data.

**Recommendations for Further Study**

1. The study should be replicated with a larger number of pupils and a wider range of subject areas represented.

2. A follow-up study should be made on the student teachers who partook in the study to determine the possible affects of their second pupil evaluations.

3. A larger control group should be given only the second evaluation at the end of a group student teachers' experiences.

4. A comparative study needs to be made with regular
teachers to determine the affects of early pupil evaluation of their teaching.

5. Other pupil evaluation forms should be used in conjunction with the one developed for this study to determine if the communication form has cross-validity.

6. The pupil evaluation forms need to be used with a longer interval between them to determine if positive change, as perceived by the pupils, takes longer than the interval used in this study.

7. A study needs to be made in which a second evaluation is given between the early and final pupil evaluations. One may determine if after the initial positive halo is reduced on the second evaluation, then significant positive change may be found between the second and final evaluations.
APPENDIX
TABLE 16
RESPONSES MADE BY SPEECH AND DRAMA PUPILS ON FORM I

<table>
<thead>
<tr>
<th>Positive Responses</th>
<th>M-17</th>
<th>F-17</th>
<th>M-16</th>
<th>F-16</th>
<th>M-15</th>
<th>F-15</th>
<th>M-14</th>
<th>F-14</th>
<th>Sums</th>
<th>Percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td>+3</td>
<td>11</td>
<td>18</td>
<td>16</td>
<td>16</td>
<td>15</td>
<td>17</td>
<td>17</td>
<td>21</td>
<td>134</td>
<td>67.0b</td>
</tr>
<tr>
<td>+2</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td>3</td>
<td>30</td>
<td>15.0</td>
</tr>
<tr>
<td>+1</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>10</td>
<td>30</td>
<td>5.0</td>
</tr>
<tr>
<td>0</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>26</td>
<td>13.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Negative Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3</td>
</tr>
<tr>
<td>-2</td>
</tr>
<tr>
<td>-1</td>
</tr>
<tr>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Response Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specific Skills</td>
</tr>
<tr>
<td>General Skills</td>
</tr>
<tr>
<td>Specific Content</td>
</tr>
<tr>
<td>General Content</td>
</tr>
<tr>
<td>Specific Value Judgments</td>
</tr>
<tr>
<td>General Value Judgments</td>
</tr>
<tr>
<td>Class Discipline</td>
</tr>
<tr>
<td>General Headings</td>
</tr>
</tbody>
</table>

aEach column presents the number of responses made by a sub-group of twenty-five pupils of a particular age and sex.

bPercentages calculated based on the sample size of 200 pupils.

cPercentages calculated on the total possible number of responses, 482.
TABLE 17
RESPONSES MADE BY ENGLISH PUPILS ON FORM I

<table>
<thead>
<tr>
<th>Positive Responses</th>
<th>M-17</th>
<th>F-17</th>
<th>M-16</th>
<th>F-16</th>
<th>M-15</th>
<th>F-15</th>
<th>M-14</th>
<th>F-14</th>
<th>Sums</th>
<th>Percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td>+3</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td>11</td>
<td>8</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>85</td>
<td>42.5%</td>
</tr>
<tr>
<td>+2</td>
<td>1</td>
<td>6</td>
<td>5</td>
<td>7</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>9</td>
<td>34</td>
<td>17.0%</td>
</tr>
<tr>
<td>+1</td>
<td>8</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>29</td>
<td>14.5%</td>
</tr>
<tr>
<td>0</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>3</td>
<td>12</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>52</td>
<td>26.0%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Negative Responses</th>
<th>M-17</th>
<th>F-17</th>
<th>M-16</th>
<th>F-16</th>
<th>M-15</th>
<th>F-15</th>
<th>M-14</th>
<th>F-14</th>
<th>Sums</th>
<th>Percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1.0%</td>
</tr>
<tr>
<td>-2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>11</td>
<td>5.5%</td>
</tr>
<tr>
<td>-1</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>19</td>
<td>9.5%</td>
</tr>
<tr>
<td>0</td>
<td>19</td>
<td>21</td>
<td>23</td>
<td>21</td>
<td>23</td>
<td>21</td>
<td>19</td>
<td>168</td>
<td>84.0%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Response Descriptions</th>
<th>M-17</th>
<th>F-17</th>
<th>M-16</th>
<th>F-16</th>
<th>M-15</th>
<th>F-15</th>
<th>M-14</th>
<th>F-14</th>
<th>Sums</th>
<th>Percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specific Skills</td>
<td>7</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>23</td>
<td>6.5%</td>
</tr>
<tr>
<td>General Skills</td>
<td>12</td>
<td>5</td>
<td>7</td>
<td>13</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>15</td>
<td>74</td>
<td>21.0%</td>
</tr>
<tr>
<td>Specific Content</td>
<td>13</td>
<td>11</td>
<td>6</td>
<td>14</td>
<td>8</td>
<td>4</td>
<td>24</td>
<td>7</td>
<td>77</td>
<td>21.8%</td>
</tr>
<tr>
<td>General Content</td>
<td>3</td>
<td>3</td>
<td>9</td>
<td>3</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>9</td>
<td>40</td>
<td>11.4%</td>
</tr>
<tr>
<td>Specific Value Judgments</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>27</td>
<td>7.7%</td>
</tr>
<tr>
<td>General Value Judgments</td>
<td>4</td>
<td>8</td>
<td>2</td>
<td>8</td>
<td>2</td>
<td>8</td>
<td>0</td>
<td>6</td>
<td>38</td>
<td>10.8%</td>
</tr>
<tr>
<td>Class Discipline</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>12</td>
<td>3.4%</td>
<td></td>
</tr>
<tr>
<td>General Headings</td>
<td>3</td>
<td>0</td>
<td>9</td>
<td>4</td>
<td>8</td>
<td>20</td>
<td>5</td>
<td>7</td>
<td>56</td>
<td>15.9%</td>
</tr>
</tbody>
</table>

*Each column presents the number of responses made by a sub-group of twenty-five pupils of a particular age and sex.

*Percentages calculated based on the sample size of 200 pupils.

*Percentages calculated on the total possible number of responses, 352.
**TABLE 18**

RESPONSES MADE BY SPEECH AND DRAMA PUPILS ON FORM II

<table>
<thead>
<tr>
<th>Positive Responses</th>
<th>M-17</th>
<th>F-17</th>
<th>M-16</th>
<th>F-16</th>
<th>M-15</th>
<th>F-15</th>
<th>M-14</th>
<th>F-14</th>
<th>Sums</th>
<th>Percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td>+3</td>
<td>16</td>
<td>14</td>
<td>14</td>
<td>16</td>
<td>16</td>
<td>21</td>
<td>19</td>
<td>132</td>
<td>66.0^b</td>
<td></td>
</tr>
<tr>
<td>+2</td>
<td>1</td>
<td>7</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>6</td>
<td>2</td>
<td>28</td>
<td>14.0</td>
<td></td>
</tr>
<tr>
<td>+1</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>11.0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Negative Responses</th>
<th>-3</th>
<th>0</th>
<th>2</th>
<th>2</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>7</th>
<th>3.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>18</td>
<td>9.0</td>
</tr>
<tr>
<td>-1</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>40</td>
<td>20.0</td>
</tr>
<tr>
<td>0</td>
<td>21</td>
<td>16</td>
<td>18</td>
<td>16</td>
<td>13</td>
<td>20</td>
<td>16</td>
<td>15</td>
<td>135</td>
<td>67.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Response Descriptions</th>
<th>Specific Skills</th>
<th>General Skills</th>
<th>Specific Content</th>
<th>General Content</th>
<th>Specific Value Judgments</th>
<th>General Value Judgments</th>
<th>Class Discipline</th>
<th>General Headings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
<td>5</td>
<td>8</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>5</td>
<td>9</td>
<td>10</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>9</td>
<td>11</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>9</td>
<td>11</td>
<td>10</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>10</td>
<td>11</td>
<td>16</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>7</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>21</td>
<td>19</td>
<td>15</td>
<td>14</td>
<td>2</td>
<td>13</td>
<td>6</td>
</tr>
</tbody>
</table>

^aEach column presents the number of responses made by a sub-group of twenty-five pupils of a particular sex and age.

^bPercentages calculated based on the sample size of 200 pupils.

^cPercentages calculated on the total possible number of responses, 470.
### TABLE 19

**RESPONSES MADE BY ENGLISH PUPILS ON FORM II A**

<table>
<thead>
<tr>
<th>Positive Responses</th>
<th>M-17</th>
<th>F-17</th>
<th>M-16</th>
<th>F-16</th>
<th>M-15</th>
<th>F-15</th>
<th>M-14</th>
<th>F-14</th>
<th>Sums</th>
<th>Percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td>+3</td>
<td>14</td>
<td>16</td>
<td>12</td>
<td>14</td>
<td>16</td>
<td>16</td>
<td>10</td>
<td>12</td>
<td>110</td>
<td>55.0b</td>
</tr>
<tr>
<td>+2</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>31</td>
<td>15.5</td>
</tr>
<tr>
<td>+1</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>21</td>
<td>10.5</td>
</tr>
<tr>
<td>0</td>
<td>7</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>8</td>
<td>7</td>
<td>38</td>
<td>19.0</td>
</tr>
</tbody>
</table>

**Negative Responses**

<table>
<thead>
<tr>
<th></th>
<th>M-17</th>
<th>F-17</th>
<th>M-16</th>
<th>F-16</th>
<th>M-15</th>
<th>F-15</th>
<th>M-14</th>
<th>F-14</th>
<th>Sums</th>
<th>Percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>2.5</td>
</tr>
<tr>
<td>-2</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>2.5</td>
</tr>
<tr>
<td>-1</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>6</td>
<td>27</td>
<td>13.5</td>
</tr>
<tr>
<td>0</td>
<td>20</td>
<td>21</td>
<td>23</td>
<td>19</td>
<td>20</td>
<td>19</td>
<td>23</td>
<td>18</td>
<td>163</td>
<td>81.5</td>
</tr>
</tbody>
</table>

**Response Descriptions**

<table>
<thead>
<tr>
<th>Category</th>
<th>M-17</th>
<th>F-17</th>
<th>M-16</th>
<th>F-16</th>
<th>M-15</th>
<th>F-15</th>
<th>M-14</th>
<th>F-14</th>
<th>Sums</th>
<th>Percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specific Skills</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>8</td>
<td>27</td>
<td>6.5c</td>
</tr>
<tr>
<td>General Skills</td>
<td>6</td>
<td>15</td>
<td>10</td>
<td>11</td>
<td>4</td>
<td>16</td>
<td>5</td>
<td>4</td>
<td>71</td>
<td>17.4</td>
</tr>
<tr>
<td>Specific Content</td>
<td>6</td>
<td>17</td>
<td>14</td>
<td>18</td>
<td>30</td>
<td>21</td>
<td>25</td>
<td>22</td>
<td>153</td>
<td>37.0</td>
</tr>
<tr>
<td>General Content</td>
<td>6</td>
<td>5</td>
<td>8</td>
<td>6</td>
<td>1</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>36</td>
<td>8.7</td>
</tr>
<tr>
<td>Specific Value Judgments</td>
<td>13</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>27</td>
<td>6.5</td>
</tr>
<tr>
<td>General Value Judgments</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>14</td>
<td>3.4</td>
</tr>
<tr>
<td>Class Discipline</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>8</td>
<td>11</td>
<td>1.9</td>
</tr>
<tr>
<td>General Headings</td>
<td>15</td>
<td>2</td>
<td>10</td>
<td>9</td>
<td>12</td>
<td>9</td>
<td>3</td>
<td>6</td>
<td>66</td>
<td>16.0</td>
</tr>
</tbody>
</table>

---

**Footnotes:**

- Each column presents the number of responses made by a sub-group of twenty-five pupils of a particular age and sex.
- Percentages calculated based on the sample size of 200 pupils.
- Percentages calculated on the total possible number of responses, 413
FORM I

DIRECTIONS: In order that this class be most beneficial to you, the students, it seems reasonable to ask you for your opinions on how it is being taught as well as what is being taught. Therefore will you read each of the following statements and give your opinion on how often the statement applies to this particular class. Use the key given below. Place the number of the phrase in the key which best gives your evaluation in the space before the statement you are considering.

To be most helpful, try to consider each statement separately. None of us are perfect and thus every statement will probably not be given the same value.

Feel free to give additional comments on any statement. Use the reverse side of this form and write the number of the statement and your comments.

You need not put your name on this form but please indicate your age, sex, and the type of class this is, such as English I, Chemistry, Speech, Drama I, etc. Your frank and honest opinion is wanted, so give serious thought to each individual statement.

Age ________ Sex ________ Class ________


1. Studying for this subject has been fun because it is made interesting and exciting.
2. The subject matter has been shown to be valuable because much of it can be used in daily life.
3. Questions are asked in a vocabulary that can be understood by all the students.
4. Phrases, such as OK or all right, are not necessarily often repeated.
5. Understanding of student questions or answers is evident because they are related to the on-going discussion.
6. Additional individual help is pleasantly given on any problem not immediately understandable.
7. Certain gestures bring attention to themselves and these are distracting.
8. Personal appearance and clothing worn are suitable and neat.
9. Ideas are discussed with words that can be understood by everyone.
10. The class is exciting and looked forward to by everyone.
11. Subject matter is easily understood because its learning is not rushed and it is gone into in depth.
12. Homework assignments have been made challenging and interesting and not just busy work.
13. The class is well-behaved and only limited control is necessary to keep it running smoothly.
14. A sense of humor brightens lectures and discussions.
15. It is evident that a lot is known about the topics taught.
16. Homework and in-class assignments are made quite clear.
17. Class discussions are lively, enlightening, and interesting and everyone participates.
18. Tests and grading is fair and does not favor anyone or anything.
19. The amount of work required is reasonable and necessary to get the most out of this class.
20. Student ideas are asked for and given due consideration.
21. In this class, many important ideas are being learned.
22. All that goes on or happens in this class is immediately recognized and effectively dealt with.
23. A pleasing personality is evident.
24. The classroom atmosphere is friendly.
25. A voice is used that is enjoyable to listen to.

Considering everything, the teacher does a good job of teaching this class.

Not only are your opinions about this class important but also what you have learned is important. Therefore please list three things that you have learned well

1. 
2. 
3. 

There might be some things that you have not learned very well, so list any that now comes to your mind.

1. 
2. 
3. 

ADDITIONAL COMMENTS:
FORM II

DIRECTIONS: Student opinion on a teacher and his course is important and valuable in determining what has been taught and learned. Therefore will you give your opinion on each of the statements presented below using the key provided. Place the number of the phrase which most closely expresses your opinion in the space preceding the statement you are considering. To be most objective, try to consider each statement separately. None of us are perfect or complete failures so every statement may not be given the same value. Feel free to give additional comments on any statement. Use the space at the bottom of this form or the reverse side and write the number of the statement and then your comments. You need not put your name on this form but please indicate your age, sex, and the type of class this is, such as, Speech I, English III, Chemistry, Drama I, etc. Your frank and honest opinion is wanted, so give each statement serious thought.

Age ______ Sex _______ Class ______

KEY: 1. Never really achieved 5. Did better than some teachers I have had
2. Still needs a lot of improvement 6. Did better than most teachers I have had
3. Still needs some improvement 7. Did as well as the best teacher I ever had
4. Did an average job

1. The voice used in teaching and discussing ideas was lively and pleasant to hear.
2. A friendly classroom atmosphere was developed.
3. A pleasing personality was evident.
4. Anything that happened in class was noticed and dealt with effectively.
5. An atmosphere was established where many important ideas were learned.
6. Student ideas were appreciated and used when possible.
7. The work assigned to students was reasonable and necessary for them to get the most out of this course.
8. Testing and grading were fair and did not favor anyone or anything to an extreme.
9. Everyone participated in class discussions because they were lively, enlightening and interesting.
10. Homework and in-class assignments were made quite clear.
11. A sense of humor was evident and enlivened discussions and lectures.
12. Very few discipline problems occurred and a minimum of control was necessary to keep the class running smoothly.
13. Rather than just giving busy work, the homework assignments were made challenging and interesting.
14. Learning of the subject matter was not rushed and it was explored in great depth.
15. Everyone looked forward to coming to this class.
16. The words used to discuss ideas were easily understood by everyone.
17. Personal appearance and dress were appropriate and neat.
18. Certain gestures drew attention and thus became distracting.
19. Student questions and answers were understood and effectively fitted into class discussions.
20. No trite phrases, such as OK or all right, were unnecessarily repeated.
21. The subject matter was made meaningful and related to current ideas and events.
22. This subject has been made exciting and interesting.

23. Considering everything, the teacher did a satisfactory job of teaching this class.

Not only are your opinions about this class important but also what you have learned is important. Therefore please list at least three things that you have learned well.

1. 
2. 
3. 

There might have been some things that you have not learned so very well, so list any that now come to your mind:

1. 
2. 
3. 

ADDITIONAL COMMENTS:
<table>
<thead>
<tr>
<th>Date Evaluated Administered</th>
<th>School</th>
</tr>
</thead>
<tbody>
<tr>
<td>Course ___________________</td>
<td>Form</td>
</tr>
<tr>
<td>Number in Class ___________</td>
<td>Teacher</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**KEY:**
- Self-evaluation
- Pupil Average
- Cooperating T.
FORM IV

Evaluator __________________ Date __________ School __________
Student Teacher ____________ Class __________

Directions: In order for a student teacher to gain the most from his first experience in teaching, it is helpful for him to receive periodic evaluation from his supervisors. On this form you are asked to respond to each of the twenty-six statements using the key given below. Consider each statement separately. Usually no student teacher will receive very high ratings on all of the statements or, on the other hand, very low ones. Hopefully this evaluation will help the student teacher pinpoint both his strengths and his weaknesses.

On the reverse side of this form, please feel free to write any additional comments that you feel have not been adequately developed and would be beneficial in providing a complete evaluation of the student teacher.

Place the number of the adjective in the key, in the space preceding each statement, which best fits your opinion of the occurrence of this behavior.


1. ___ Because of planning and the procurement of materials, the study of this subject is being made interesting and exciting for the pupils.
2. ___ In discussion and lectures the subject matter is made meaningful because it is related to student experiences and current events.
3. ___ Students are asked questions in a vocabulary that is appropriate and understandable to them.
4. ___ Trite expressions are absent or seldom heard in class discussions or lectures.
5. ___ Every attempt is made to integrate student responses into the ongoing class discussion.
6. ___ Additional individual help is pleasantly given on any idea or problem that is not immediately understood.
7. ___ Certain gestures or mannerisms are distracting.
8. ___ Dress and personal appearance is satisfactory and consistent.
9. ___ Explanation and elaboration of ideas are presented with a vocabulary well-adapted to the students' understanding.
10. ___ Enthusiasm for learning and class morale is high and the class seems to enjoy coming to class.
11. ___ Study of the subject matter is not being rushed and there is evidence of in-depth probing for understanding of the course content.
12. ___ Homework assignments are a logical extension of in-class work and not just busy-work.
13. ___ Student self-discipline is fostered so that there is a minimum need for teacher disciplining of students.
14. ___ A sense of humor enlivens discussions and reduces tensions in a tense situation.
15. ___ The teacher is knowledgeable about the subject matter.
16. ___ Explanation of in-class and homework assignments is quite clear.
17. ___ Class discussion are lively and the teacher encourages everyone to participate in them.

Please turn over
<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>18.</td>
<td>Testing and grading are fair and do not favor anything or anyone.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19.</td>
<td>Work assigned to students is not excessive but necessary for adequate learning of the subject matter.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20.</td>
<td>Student ideas are sought and acted upon when there is consensus among the students and the teacher.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21.</td>
<td>The importance of learning the subject matter is well-developed.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22.</td>
<td>The teacher is perceptive of each individual student's behavior and, when it becomes distracting, deals with it effectively.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23.</td>
<td>The personality of the teacher is such that students enjoy being around him.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24.</td>
<td>The teacher is friendly with everyone with whom he works.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25.</td>
<td>The voice of the teacher is easy to listen to and expressive in communicating ideas.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26.</td>
<td>Considering everything, the teacher does a satisfactory job of teaching.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**ADDITIONAL COMMENTS:**


BIBLIOGRAPHY

Books


Periodicals


Bryan, Roy C. "Benefits Reported by Teachers Who Obtained Written Student Reactions," Educational Administration and Supervision, 28 (January, 1942), pp. 69-75.


Knower, Franklin H. "A Model for a Communicology," 


Tuckman, B.W., and Oliver, W.F. "Effectiveness of Feedback to Teachers as a Function of Source," *Journal of Educational Psychology*, 59 (August, 1968), pp. 297-301.


Reports


Unpublished Material