Increasing the cognitive level of classroom questions: an application of Bloom's Taxonomy of Educational Objectives.

George Thomas Farley

University of Massachusetts Amherst

Follow this and additional works at: https://scholarworks.umass.edu/dissertations_1

Recommended Citation

https://scholarworks.umass.edu/dissertations_1/4182

This Open Access Dissertation is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Doctoral Dissertations 1896 - February 2014 by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
INCREASING THE COGNITIVE LEVEL OF CLASSROOM QUESTIONS: AN APPLICATION OF BLOOM'S TAXONOMY OF EDUCATIONAL OBJECTIVES

A Dissertation Presented

by

George T. Farley
Fellow in Applied Educational Research

Submitted to the Graduate School of the University of Massachusetts in partial fulfillment of the requirements for the degree of

DOCTOR OF EDUCATION

May 1, 1968

Major Subject: Administration
Copyright by
George T. Farley
1968
All Rights Reserved
INCREASING THE COGNITIVE LEVEL OF CLASSROOM QUESTIONS: AN APPLICATION OF BLOOM'S TAXONOMY OF EDUCATIONAL OBJECTIVES

A Dissertation

by

George T. Farley
Fellow in Applied Educational Research

Approved as to style and content by:

(Chairman of Committee)

(Head of Department)

(Member)

(Member)

(Member)

(Month) (Year)
ACKNOWLEDGMENTS

In writing this dissertation, the writer is indebted to a number of persons. To Dr. Ambrose A. Clegg Jr., who, over a period of four years gave freely of his advice and encouragement, my heartfelt thanks. To Dr. William C. Wolf and Dr. Ovid F. Parody a special word of appreciation for their guidance and scholarly evaluations. To Dr. Emma M. Cappelluzzo and Dr. Harry Schumer a word of thanks for their willingness to accept the arduous task of readers.

My thanks are also extended to Dr. Bernard J. Fleury, Dr. Gerald H. Lunney, and Mr. Paul A. Leary who contributed much during the experimental phase of this study. I owe a special debt of gratitude to Dr. Fleury whose personal association and kind assistance played a major role in the successful completion of this study.

Finally, to my wife, Marge, a simple thanks seems so inadequate. Without her unselfish devotion and sacrifice, none of this would have been possible.

George T. Farley
TABLE OF CONTENTS

LIST OF CHARTS AND FIGURES vii
LIST OF TABLES viii

Chapter

I. THE PROBLEM ......................................... 1

Statement of the Problem
Background
Purpose of the Study
Significance of the Problem
Definition of Terms
Limitations of the Study

II. RELATED RESEARCH ................................. 16

Early Studies
Social Climate and Social Interaction
Teacher Characteristics and Teaching Patterns
Cognitive Interaction
Conclusions

III. METHODS AND MATERIALS ....................... 37

Subjects
Investigators
Setting
Experimental Procedures
Unit of Measurement
Rules for Scoring the TPQI
Analysis of Data

IV. PRESENTATION AND ANALYSIS OF DATA ....... 53

Step One: Tests of Hypotheses One and Two
Step Two: Tests of Hypotheses Three, Four and Five
Step Three: Test of Hypothesis Six
Step Four: Test of Hypothesis Seven
Step Five: Test of Hypotheses Eight and Nine
V. CONCLUSIONS .............................................. 75

Discussion of Hypothesis One
Discussion of Hypothesis Two
Discussion of Hypothesis Three
Discussion of Secondary Hypotheses
General Conclusions
Recommendations

VI. SUMMARY .................................................. 91

Statement of Problem
Related Research
Objectives
Procedures
Design
Findings
Conclusions
Recommendations

APPENDIX .................................................... 101

The Bloom-Sanders' Taxonomy of Educational Objectives: Cognitive Domain
The Flanders-Amidon Model of Categories for Interaction Analysis
Weekly Procedures
Teacher-Pupil Question Inventory (TPQI) Composite Scores
Individual TPQI Scores

BIBLIOGRAPHY ............................................... 117
# LIST OF CHARTS AND FIGURES

<table>
<thead>
<tr>
<th>Chart/Medium</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chart</td>
<td>Study Synopsis of the Student-Teacher Experiment</td>
<td>8</td>
</tr>
<tr>
<td>Chart</td>
<td>Study Synopsis of the Co-operating Teachers Experiment</td>
<td>10</td>
</tr>
<tr>
<td>Figure</td>
<td>The Flanders-Amidon Matrix</td>
<td>22</td>
</tr>
<tr>
<td>Figure</td>
<td>An Example of Taba's Flow Chart for Raising Cognitive Behavior</td>
<td>27</td>
</tr>
<tr>
<td>Figure</td>
<td>Design Associated with Hypotheses One and Two</td>
<td>49</td>
</tr>
<tr>
<td>Figure</td>
<td>Design Associated with Hypotheses Three through Seven</td>
<td>50</td>
</tr>
<tr>
<td>Figure</td>
<td>Frequency and Percent of Questions Asked at the Different Cognitive Levels</td>
<td>51</td>
</tr>
<tr>
<td>Figure</td>
<td>Student-Teachers x Trial Interaction</td>
<td>62</td>
</tr>
<tr>
<td>Figure</td>
<td>Co-operating Teachers x Trial Interaction</td>
<td>62</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table | Page
--- | ---
1. Investigators' (R1-R3) Composite Scores for Experimental (T1) and Control Group (T2) Over Three Trial Periods (C1-C3) | 56
2. Summary of Analysis of Variance Test of Mean Scores for Treatment and Investigators | 57
3. Co-operating Teachers' (R4-R9) Composite Scores for Experimental Group (T1) Over Three Trial Periods (C1-C3) | 59
4. Summary of Analysis of Variance Test of Mean Scores for Co-operating Teachers, Student-Teachers, and Trials | 60
5. Raters' (R1-R9) Composite Scores for Experimental Group (T1) Over Three Trial Periods (C1-C3) | 63
6. Summary of Analysis of Variance Test of Mean Scores for Raters (R1-R9) | 64
7. Co-operating Teachers Variances for the Three Trial Periods | 66
8. Frequency and Percentage of Student-Teachers' Questions Arranged by Cognitive Levels | 68
9. Chi-square Tests Comparing Experimental Group with (1) Control Group and (2) Davis-Tinsley Results | 69
10. Frequency and Percentage of Combined Student-Teacher Questions and Pupil Responses Arranged by Cognitive Levels | 71
11. Comparison of Totals of Frequency and Percentage Teacher Questions and Combined Teacher Questions and Pupil Responses Arranged by Cognitive Levels (From Tables 8 and 10) | 72
CHAPTER I

INTRODUCTION

Statement of the Problem

In education today, high-level goals are often stated as objectives, but these goals are seldom achieved. Scannell and Stellwagen (1960) and Tyler and Okumu (1965) found that rarely was there a direct relationship between the levels of stated goals and levels of performance. Inn (1966) concluded that when teachers seek varied experiences for their pupils, little concern is given to the conceptual objective to be developed.

Educators use such statements as, "developing critical thinking," "leading students to the discovery of knowledge," and "making students inquirers" - all of which imply that students should do more than master a plethora of facts. Yet, in reviewing the studies aimed at determining the cognitive level attained in the majority of classrooms, one finds that most time is given over to memory-type learning. Davis and Hunkins (1966) found this to be true in examining text-book questions. Approximately eighty-five percent of all text-book questions were at the memory level. They projected that the level of text-book
questions greatly influences the overall cognitive behavior prevalent in the classroom. Davis and Tinsley (1967) found that, when they applied a modified version of Bloom's Taxonomy of Educational Objectives to interpreting the level of classroom dialogue, approximately eighty-five percent of the questions and responses were either at the memory or comprehension level. Gallagher (1965) had similar results when he applied Guilford's model to classroom behavior.

Fenton (1966:19) felt that our failure to reach high-level objectives is because the objectives are "imprecisely stated" and are too diffuse to be dealt with accurately. In other words, we must learn to clearly define our objectives. Sanders (1966) offered a similar view when he said that critical thinking is usually ill-defined and thus the various thought processes included in this concept are lost because of the vague understanding of the phrase. Taba and Elzey (1964) gave two reasons for this state of affairs: (1) Thinking has been treated as a global process. (They felt that it should be considered from its specific elements). (2) Heavy reliance has been placed on two questionable assumptions: (a) reflective thinking cannot take place until a sufficient body of factual information has been accumulated, and (b) thought is an automatic by-product of studying certain subjects.
The nature of the problem appears to evoke questions related both to the pre-service and in-service training of teachers. The following questions, which are specific aspects of the general problem, were asked by this study:

(a) How can student-teachers be taught to operate, within the classroom, at a higher level of cognitive behavior than is achieved at present?

(b) How can regular classroom teachers be brought to recognize and agree upon the cognitive levels of behavior that occur during a given classroom lesson?

**Background**

Efforts to aid both the regular teacher and the student teacher to raise the level of classroom effectiveness can be traced back fifty years to the work of Horn (1914). His approach was to tabulate the number of responses emitted by each student. Puckett (1928) took up and expanded Horn's work. Puckett's contribution consisted of developing a set of symbols, placed next to the student's name, indicating the source of stimulus and quality of answer proffered by the student.

The study of teacher responses was introduced by Wrightstone (1934). In some ways Wrightstone's method resembles Puckett's in that symbols were placed alongside the student's name. These symbols were number-letter
combinations and referred to teacher behavior as the teacher interacted with the students.

The measurement of classroom social climate began with the Thomas Study (1929). This study can be classified under the broader heading of sociometry which had its beginning around 1925 (Moreno, 1956). Thomas focused on the observable social interactions among students. The works of Anderson (1945), Withall (1949), Medley-Mitzel (1959), and Flanders (1960) were stimulated by the findings of Thomas.

Another approach to improving teacher effectiveness was based on the preconception that certain personal characteristics will promote good teaching behavior. The monumental studies of Barr (1929, 1961), and Ryans (1960) attest to the large number of educators who, through the years, have felt that personal characteristics are vitally important to raising the level of classroom performance. After years of investigations, Barr and Ryans were able to conclude only that certain patterns of affective qualities such as "warmth" and "understanding" were characteristics of good teaching traits.

None of the studies mentioned above has dealt directly with cognitive behavior, however. Smith (1959) and Wright (1959) have led the way in this area. Taba (1965), Taba and Elzey (1964), and Taba et al. (1964) also dealt with this problem when they attempted to teach
certain cognitive skills in the classroom. Recently, Davis and Tinsley (1967) and Clegg et al. (1967a, 1967b) reported studies in which they attempted to determine the level of pupil-teacher cognitive interaction. They devised a measuring instrument incorporating the levels of cognitive behavior based on Bloom's Taxonomy of Educational Objectives.

This present study continued investigation into the area of cognitive interaction. It employed some of the techniques used in the Davis-Tinsley and the Clegg et al. studies. A fundamental difference, however, was that a control group was compared to an experimental group.

Purpose of the Study

Objectives

The three main purposes of this study were to determine if: (1) student-teachers who had received instruction in the use of Bloom's Taxonomy would operate, within the classroom, at a higher cognitive level than student-teachers who had not received instruction in the Taxonomy (see Appendix A, p. 102 for a definition of the categories making up the Taxonomy); (2) trained observers could correctly identify the cognitive classroom behavior level stimulated by the classroom instructor; (3) regular teachers, as observers, could correctly identify the cognitive classroom behavior level stimulated by the class-
room instructor.

The above objectives, when written in behavioral terms, read as follows:

(1) Student-teachers will demonstrate their ability to raise the cognitive level of classroom performance by conducting social studies lessons which include a significantly higher ($p < .01$) proportion of above-memory questions and responses than achieved by a control group of student-teachers. To determine the cognitive level of performance, Bloom's Taxonomy will be used to evaluate each question and response.

(2) Trained observers will demonstrate their ability to correctly identify the levels of cognitive classroom behavior created by the student-teacher by using a rating form (based on Bloom's Taxonomy) to evaluate the lessons. Statistical agreement ($p < .01$) of rating scores of trained observers will be used as a minimum criterion.

(3) Co-operating teachers will demonstrate their ability to correctly identify the levels of cognitive classroom behavior created by student-teachers by using a rating form (based on Bloom's Taxonomy) to evaluate the lessons. Statistical agreement ($p < .01$) of rating scores of all co-operating teachers will be used as a minimum
Variables

The independent variable was knowledge of Bloom's Taxonomy and application of it to a classroom instructional situation. The Taxonomy was introduced to student-teachers in the experimental group as they began their student-teaching assignments. Weekly instruction in the Taxonomy continued through the eight-week period of student-teaching. (See Appendix C, p. 106 for the weekly supervisory schedule related to the experimental and control groups).

The dependent variable was the score on the Teacher Pupil Question Inventory (TPQI) similar to that used by Clegg et al. (1967a, 1967b). Student-teachers making up the experimental group were compared with student teachers making up the control group (see Appendix D, p. 109). The chart below provides a study synopsis for the student-teacher experiment.
CHART 1

STUDY SYNOPSIS OF THE STUDENT-TEACHER EXPERIMENT

<table>
<thead>
<tr>
<th>Treatments</th>
<th>Modus Operandi</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>#1 Experimental Group</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Independent Variable =</strong> knowledge of Bloom's Taxonomy and its application in classroom situations.</td>
<td><strong>Instrumentation</strong></td>
</tr>
<tr>
<td>Treatment of student-teachers consisted of (1) instruction in applying Bloom's Taxonomy to classroom teaching procedures and (2) self-evaluation of classroom lessons. The latter was achieved by having the student-teacher listen to taped recordings of lessons which she herself had taught and evaluating them according to Bloom's Taxonomy.</td>
<td>The TPQI was based on Bloom's Taxonomy and was used to determine if either the control or experimental group achieved at a higher cognitive level.</td>
</tr>
</tbody>
</table>

**#2 Control Group**

No instruction in Bloom's Taxonomy was afforded the student-teachers in this group but equivalent time was given to them.

Interaction analysis of classroom lessons was performed here by having the student-teachers listen to their own taped lessons. The Flanders-Amidon model was used to create a placebo effect.

**Procedure**
The students in each group were evaluated by three investigators who appear in the design as raters one, two, and three. ($R_1$, $R_2$, $R_3$).

Three tapes of each student-teacher were rated. These tapes were made during the third, fifth, and seventh weeks of the teacher-training period.
The variables related to the co-operating teachers were similar to those of the student-teachers. Their application was slightly different, however. The treatment variable, which was knowledge and application of the Taxonomy, also included the observation of student-teachers by co-operating teachers. The co-operating teachers were first given practice in rating student-teachers and these rating scores were discussed and compared with rating scores of the investigator (R₁).

In the design, the co-operating teachers appear as raters four through nine (R₄ - R₉). Each of the six co-operating teachers rated separately each student-teacher's taped lesson. The purpose of comparing co-operating teachers' ratings was to determine if a group of experienced teachers could agree in their evaluation of an individual student-teacher's performance.

The dependent variable again was the TPQI, but it focused on how the co-operating teachers agreed with one another. This differed from the way the TPQI was used with the student-teachers where the emphasis was on the cognitive level that was achieved. The chart below provides a study synopsis of the co-operating teachers experiment.
CHART 2

STUDY SYNOPSIS OF THE CO-OPERATING TEACHERS EXPERIMENT

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Modus Operandi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent Variable = knowledge of Bloom's Taxonomy and its application in evaluating lessons taught by student-teachers.</td>
<td>Dependent Variable = Score on Teacher Pupil Question Inventory (TPQI)</td>
</tr>
</tbody>
</table>

The six co-operating teachers made up the experimental group. There was no control group. Co-operating teacher ratings were compared to determine if agreement of raters would be achieved.

Procedure

Students of the experimental group only were evaluated by the six co-operating teachers who appear in the design as raters four through nine (R4-R9).

Three tapes of each student-teacher were rated. These tapes were made during the third, fifth, and seventh weeks of the teacher-training period.

Hypotheses

The three main hypotheses, stated in null form, were:

1. There is no difference in the level of cognitive classroom behavior achieved by a group of student-teachers who have had instruction in the use of Bloom's Taxonomy and a group of student-teachers who have had no instruction in the use of the Taxonomy.

2. There is no difference among investigators
(R₁ - R₃) in their rating of student-teachers, using the TPQI as a measuring device.

(3) There is no difference among co-operating teachers (R₄ - R₉) in their rating of student-teachers using the TPQI as a measuring device.

Within the context of the above hypotheses and the method proposed to obtain results, it was possible also to test several secondary hypotheses:

(4) There is no difference in the level of cognitive classroom behavior achieved by student-teachers who have been instructed in the use of Taxonomy.

(5) There is no increase over time in the level of cognitive classroom behavior of student-teachers who have been instructed in the use of the Taxonomy.

(6) There is no difference between the rating scores of the investigators (R₁ - R₃) and the rating scores of the co-operating teachers (R₄ - R₉).

(7) There is no increase in rating agreement of co-operating teachers over time.

(8) There is no difference in the percentage of memory questions asked by a group of student-teachers who have had training in the use of the Taxonomy (experimental group) and a group of student-teachers who have had no training in the
use of the Taxonomy (control group).

(9) There is no difference in the percentage of memory questions asked by a group of student-teachers who have had training in the use of the Taxonomy (experimental group) and a group of student-teachers, previously reported in the literature, who had no training in the use of the Taxonomy (Davis and Tinsley, 1967).

**Significance of the Problem**

The problem is important because it is intrinsically related to teacher improvement, teacher evaluation, and ultimately to pupil achievement. It is concerned with student-teacher training and in-service training of experienced teachers. It centers on classroom dialogue and the technique of asking questions. The measuring instrument is so designed that it records verbal behavior as it happens.

This present study was designed to aid both the student-teacher and the experienced teacher. It concentrated on classroom dialogue, using a method which, hopefully, would improve the cognitive interchange among teacher and students. The importance of classroom dialogue cannot be overstated. Withall (1949) and Flanders (1960) claimed that a teacher's verbal behavior is a representative sample of his total behavior. Flanders further
reported that about two-thirds of the time someone is talking in the classroom. It seems important, then, that classroom dialogue should be meaningful. One finds, however, that research into methods for improving the cognitive level of classroom dialogue is quite meager. Aschner (1961) and Taba and Elzey (1964) and Taba (1966) have developed strategies for raising cognitive classroom behavior. Clegg et al. (1967a, 1967b) suggested that there are verbal cues that need to be known by both teachers and students in order to raise the cognitive level of classroom questioning. Yet, work in this area has only just begun. Not only must different theories be tested, but devices must be constructed to determine the value of these new approaches.

The measuring instrument used in this study was quite simple and regular classroom teachers were able to employ it after a few practice sessions. It had the advantage of being able to record cognitive behavior as it occurred. For years, educators have expressed their dissatisfaction with those devices which were used either before or after the lesson was presented. Puckett (1928) and Flanders claimed that such instruments were more susceptible to observer bias than those which were used as the lesson was being taught.
Definition of Terms

Cognitive Behavior refers to the intellectual process involved in any thinking act. According to Bloom (1956), the process may reach different levels of complexity with each level including all the levels below it. Recent studies by Kropp and Stoker (1966) tend to confirm the hierarchal nature of the Taxonomy as originally postulated by Bloom. The six levels included in the Taxonomy of Educational Objectives are: (1) knowledge (memory), (2) comprehension, (3) application, (4) analysis, (5) synthesis, and (6) evaluation.

Cognitive Interaction refers to the level of intellectual discourse that is taking place among individuals. In this study it refers specifically to verbal questions and responses between teacher and pupils and between pupil and pupil.

Limitations of the Study

Comparatively small samples of twelve student-teachers and six co-operating teachers were involved in this study. Thus, although both groups may be considered representative of their populations, generalizations must be limited due to sample size. A second limitation is related to the dependent variable. The instrument that measured results (TPQI) determined the cognitive behavior level achieved in the classroom. No device was employed
to measure pupil achievement, however, as standardized tests are not available to measure the more sophisticated levels of the Taxonomy. The approach taken here seems justified, though, because of Hunkins' findings. Hunkins (1967) reported that when the cognitive level within the classroom was raised, pupil achievement was increased. This present study assumed, therefore, that pupil achievement would increase as cognitive classroom behavior was raised. Consequently, this study concentrated on raising the cognitive behavior in the classroom and used a measuring device to determine if the cognitive level had actually been raised.
CHAPTER II

RELATED RESEARCH

In discussing the different attempts at increasing teacher effectiveness, this investigator has classified them under four headings: early studies, social interaction, teacher characteristics, and cognitive interaction. Briefly, early studies reviewed the studies that were directed at analyzing separately either student statements or teacher statements for supervisory purposes. Social interaction was concerned with pupil-teacher social dialogue and its effect on classroom climate. Teacher characteristics discussed the qualities which are supposed to be possessed by good teachers. Cognitive interaction considered the different levels of thinking and how cognitive classroom behavior might be improved.

Early Studies

Before World War I, educators showed an interest in developing objective measures of teacher behavior for supervisory purposes. Horn (1914) proposed a method whereby either a circle or a square was placed beside the student's name. The circle indicated a verbal response of the student. The square represented a response other
than verbal, such as, writing on the blackboard. Horn's main objective was to determine how well the teacher achieved a distribution of responses among his students. Although this method succeeded in tabulating the number of responses emitted by each student, it did not provide a means for determining the quality of each response.

Puckett (1928) developed a method patterned after that of Horn. He tabulated students' questions as well as their responses. His symbols also indicated the quality of the response which the student had given. This method differed from most checklists of the day in that the rater was engaged in recording the symbols immediately after the responses were made. Puckett contended that other checklists, which were filled in after the class period was over, contained only impressions of what went on in the classroom. Puckett admitted, however, that his instrument became subjective at the point where the rater was forced to determine the quality of the response.

Wrightstone (1934) developed a number-letter code. These symbols were also recorded next to the student's name. The main difference between this work and that of Puckett lies in the fact that Wrightstone considered teacher responses to individual students rather than student responses to teachers. Although his rating system was somewhat complicated, Wrightstone developed a
More sophisticated scoring method for analyzing results.

Social Climate and Social Interaction

The measurement of class climate began with the work of Dorothy Thomas (1929). She felt that existing rating forms and case studies did not take into account the tremendous complexity of any behavioral act. To obtain a more complete understanding of the classroom situation, one must record the interactions of pupils with pupils and pupils with teachers. As stated earlier, this approach was part of the sociometry movement identified with the work of Moreno (1956). Thomas' investigations included: (1) observing individual children and their actual movements around the room, (2) recording every physical contact made by a child and the kind of response that ensued, (3) tabulating a child's verbal behavior as well as everything said to him. This present study was interested only in the third point as it was restricted to investigating classroom questions.

Working on one phase of the Thomas study, Fisher (1932) analyzed transcripts of classroom dialogue. Fisher was interested in the type of language used by children related to grammatical form, types of sentences used, and percentage of the use of self-referents and other-referents. Fisher was able to conclude only that there was correlation among items which could be identified
with egocentricity.

Anderson (1945) took up the work of Thomas. He developed two main categories related to social interaction. They were: (1) dominant behavior, and (2) integrative behavior. These were first used to classify children's acts. Later, they were applied to teacher behavior. Subsequently, intensive work was done to subdivide these categories into several parts. Anderson was successful in identifying teachers who exhibited both dominant and integrative behavior. Subsequent studies by Withall (1949) and Mitzel and Rabinowitz (1953) supported Anderson's findings.

Withall found high correlation between his scale and Anderson's scale. Although Withall developed seven categories\(^1\), he considered the first three under the heading of "problem-centeredness" and the last three under "teacher-centeredness." Withall equated problem-centeredness with integrative behavior and teacher-centeredness with dominant behavior.

Mitzel and Rabinowitz found that, in using Withall's

---

1. Withall's scale is a continuum, moving from what he termed "problem-centeredness" \((1, 2, 3)\) to "teacher-centeredness" \((5, 6, 7)\). The categories are: 1. Learner-supportive, 2. Acceptant and clarifying, 3. Problem-structuring, 4. Neutral statements, 5. Directive or hortative, 6. Reproving and deprecating, 7. Teacher self-supporting.
scale to determine classroom climate, agreement of raters could be achieved. They also projected a very important hypothesis for future investigations, namely, that the consistency of climate may be more important for learning than the type of teacher behavior. They suggested that consistency reduces student anxiety, which in turn, facilitates social and intellectual growth.

Several years later, Medley and Mitzel (1958) developed an instrument for observing classroom climate known as the Observation Schedule and Record (OScAR). Half of the OScAR is based directly on Withall's model. The Expressive Behavior Section, as this half is called, is made up of five of Withall's categories. Medley and Mitzel eliminated the category headings entitled "Acceptant and Clarifying" and "Teacher Self-Supportive."

Within the Expressive Behavior Section, space is also provided for the observer to interpret facial expressions and gestures. These are recorded simply as approving or hostile.

Another sophisticated device developed for measuring classroom climate is Flanders' Interaction Analysis Matrix (1960). The Flanders model employs ten categories - seven are related to teacher talk, two to student talk, and one to "other" situations. A 10 x 10 matrix is used to tabulate responses. (See Figure 1.)
The matrix is analyzed to determine in what areas there appear to be large proportions of tally marks. Certain areas indicate a particular type of classroom climate. The figure below illustrates how Flanders interprets his matrix. For example, if a large number of tallies is found in the 1-3 x 1-3 area, the teacher is considered to exert indirect influence upon his students.

According to Medley and Mitzel (1963) Flanders' scheme for analyzing results is simple and extremely ingenious.

In this present study, the Flanders Interaction Analysis method was introduced to the student-teachers of the control group in a similar manner as the Taxonomy was introduced to the student-teachers of the experimental group. It was planned, therefore, that the control group would be considered a placebo and that any Hawthorne Effect would have no greater influence upon the experimental group than upon the control group.

**Teacher Characteristics**

At the time that Thomas was observing Classroom Social Climate, Barr (1929) was beginning to analyze teacher characteristics. Barr's theory was that if people are appointed to teaching positions who seem to possess
FIGURE 1
THE FLANDERS-AMIDON MATRIX

<table>
<thead>
<tr>
<th>Category</th>
<th>Second Three Seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5 6 7 8 9 10</td>
</tr>
<tr>
<td>1. T. accepts feelings</td>
<td>Indirect</td>
</tr>
<tr>
<td>2. T. praises/encourages</td>
<td>Influence</td>
</tr>
<tr>
<td>3. T. accepts/uses ideas</td>
<td></td>
</tr>
<tr>
<td>4. T. asks questions</td>
<td></td>
</tr>
<tr>
<td>5. T. gives lecture</td>
<td></td>
</tr>
<tr>
<td>6. T. gives directions</td>
<td>Direct</td>
</tr>
<tr>
<td>7. T. criticizes/justifies</td>
<td>Influence</td>
</tr>
<tr>
<td>8. St. responds</td>
<td></td>
</tr>
<tr>
<td>9. St. initiates</td>
<td></td>
</tr>
<tr>
<td>10. Silence/confusion</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td>A. Teacher Talk</td>
</tr>
<tr>
<td></td>
<td>B. Student Talk</td>
</tr>
</tbody>
</table>

2. Adapted from Flanders and Amidon (1963)
the qualities that enhance the learning situation, then teacher effectiveness will be improved. For the next thirty years, Barr and many of his students investigated different aspects of this problem. Barr (1961) conceded, however, that most of the work was descriptive in character and that the investigator could make no attempt to generalize.

The Ryan studies stand today as the outstanding research in the area of teacher traits and characteristics. Funded by the Grant Foundation in 1948, the studies took over ten years to complete. One of the objectives of this ambitious project was to "identify and analyze patterns of classroom behavior, attitudes, viewpoints and intellectual and emotional qualities which may characterize teachers" (1960:9). The studies involved over one hundred separate research projects which analyzed classroom behavior, surveyed teacher characteristics, and studied special areas of teacher behavior.

Through factor analysis, Ryan found that certain traits were associated with good teachers. They were: "warm, systematic, stimulating, imaginative, and surgent" (1960:388). Bowers and Soar (1961) came to much the same conclusions although their focus was on the effects of a course in human relations. Bowers and Soar found that teachers receiving such training were more accepting and
permissive in their association with pupils and more democratic in their idea of educational leadership.

Peronto (1961) found various combinations of personality traits which he identified with good and poor teachers. No single ideal type of teacher could be identified, however. Peronto concluded that knowledge of subject matter and understanding of pupils combined with professional knowledge are the only definite discriminating agents thus far established to distinguish between good and poor teachers.

Turning from teacher personality traits to teacher classroom behavior, some broad findings can be reported. Stevens (1912), Barr, Flanders, and Giammateo (1963) found that approximately sixty-five percent of the talking in the classroom was done by the teacher. Jayne (1945) found it to be closer to forty-six percent. All concurred that good teachers do less talking than poor teachers. Flanders, Amidon and Flanders, and Schantz (1963) found that teachers who exert indirect influence (according to the Flanders model) were more successful in raising pupil achievement.

**Cognitive Interaction**

Within the last ten years, some educators have turned their attention to analyzing classroom cognitive
dialogue. Smith (1959) and Wright (1959) are generally credited as the pioneers in this area. Smith classified classroom verbal behavior either as an episode or a monologue. His study dealt mostly with the episode which is defined as a verbal interaction among teacher and pupils or pupils and pupils. According to Smith, there are three parts or phases to an episode: (1) the entry phase, (2) the continuing phase, and (3) the closing phase. Smith concentrated on the entry phase. He categorized the different opening statements according to the types of responses they were supposed to elicit. Although Smith identified thirteen categories of cognitive skills, one can perceive similarities between them and those found in Bloom's Taxonomy. For example, Smith labeled one category as "Comparing and Contrasting," while Bloom categorized his as "Analysis." Both men expect the student, in performing this mental operation, to break down the material into its parts.

Wright investigated classroom dialogue from three dimensions: (1) ability to think, (2) appreciation of mathematics, and (3) curiosity and initiative. She also developed an instrument for measuring classroom verbal interaction. This instrument was designed specifically for use in mathematics classes, and Wright has stated that in order to obtain satisfactory results, the
observer must have a strong math background. Wright felt, however, that with some modifications the instrument could have broad application.

Taba (1965), Taba and Elzey (1964), and Taba et al. (1964) reported on studies designed to develop the thought process in the classroom. Their basic assumption was that the thought process follows certain sequences. Their findings indicated that basic patterns of the cognitive process are identifiable and consist of three main tasks: (1) concept formation, (2) interpretation of data, and (3) application of principles. They maintained that it is important for the teacher to know at what stage a student is in order for him to obtain mastery. They also felt that the material must be presented in the correct mode. (They refer here to Piaget's interpretation of the term "mode." It is the manner in which a person obtains information). Piaget claims there are three broad stages or modes of human development, namely, pre-operational, concrete-operational, and hypothetical-deductive (Piaget, 1964; Flavell, 1963). Thought process and mode were the variables with which Taba and her associates worked.

Taba et al. also considered the level of questioning. They identified four levels: (1) focusing (F), (2) extending thought on the same level (X), (3) lifting thought to a higher level (L), and (4) controlling (C).
Flow charts were drawn to illustrate cognitive behavior and how the level of thought was raised. An example is shown below:

**FIGURE 2**

AN EXAMPLE OF TABA'S FLOW CHART FOR RAISING COGNITIVE BEHAVIOR

Klebaner developed a set of guidelines for asking good questions. The guidelines focused on the following headings: timing, justification, clarity, formulation, appropriateness, flexibility, answers, logic, sequence, and thought. Klebaner insisted that a depth approach to classroom questioning is necessary in order to achieve any good, clear line of thinking. She also submitted that pupils should come to realize that each question has a specific purpose for them and this purpose "must be perceived in order to answer properly" (1967, 77). This statement seems to suggest that if students are taught the different levels of intellectual thought (such as
outlined by Guilford or Bloom and discussed later in this chapter), they will be better equipped to answer questions at the desired cognitive levels of teacher expectancy.

Bellack et al. (1966) concluded that teacher-pupil interaction leads to certain cyclical patterns which they termed "teaching cycles." From the standpoint of content, they identified four types of teaching cycles, namely: (1) substantive with associated, (2) substantive-logical meanings, (3) instructional with associated, and (4) instructional-logical meanings. The substantive-logical meanings cycle is synonymous with cognitive behavior. Bellack and his associates found that between sixty-one and eighty-eight percent of classroom time was devoted to the substantive-logical category. This category was subdivided into three parts:

1. Empirical meanings - fact setting and explaining
2. Analytic meanings - defining and interpreting
3. Evaluative meanings - opining and justifying

Bellack et al. found that the median time devoted to empirical meanings was six times that of analytic and evaluative meanings combined. This ratio is very similar to the results obtained by Gallagher (1965) and Davis and Tinsley (1967) whose studies are discussed later in this chapter.
Two Models for Analyzing Cognitive Behavior

The concluding section of this chapter will discuss two models which are used to classify intellectual operations. They are Guilford's Model (1959) and Bloom's Taxonomy of Educational Objectives: The Cognitive Domain (1956).

Guilford's Model

This model places all cognitive behavior into a 5 x 4 x 6 matrix. The three classifications associated with the model are: operations, content, and products. The Gallagher-Aschner studies used only the operations classification in conducting experiments related to classroom questioning. The headings classified under operations are defined below.

(1) Cognition - which is the discovery, rediscovery, or recognition of information.
(2) Memory - which is the retention of what is known.
(3) Convergent Thinking - which is the generation of new information from known and remembered information. (This leads to one right answer or to a recognized best answer.)
(4) Divergent Thinking - which is also generated from known or remembered information. (This leads to a variety of answers, however.)
(5) Evaluation - which is the reaching of decisions
Aschner (1961) stated that teachers regularly provide four main types of questions: remembering, reasoning, evaluating, and creative thinking. She maintained that a teacher's strategy should be first to design and plan the kind of thinking task to be set, and second to fit the form and phrasing of the question to the task.

In developing an instrument to measure the kinds of thinking that are prevalent in the classroom, Gallagher and Aschner (1963) employed the last four categories of Guilford's model. They found that when the percentage of divergent thinking questions from the teachers was high, the divergent thinking production from children was also high. The converse was also true.

Gallagher (1965) later reported that the majority of teacher questions and pupil responses fell into the memory area. He stated that the more a class was taught by the lecture method, the higher percentage of memory type questions and responses occurred. Convergent thinking was the second most frequent category used. The percentage of responses in the evaluative and divergent thinking categories was found to be extremely low unless the teacher made a deliberate effort to stimulate responses in these areas.
Bloom's Taxonomy of Educational Objectives

The Taxonomy was originally designed to aid curriculum makers in specifying educational objectives. It was also considered helpful to teachers in making lesson plans and achieving goals of a teaching unit. Recently, it has been used to analyze examinations (Pfeiffer and Davis, 1965), evaluate textbooks (Davis and Hunkins, 1965), construct oral and written questions (Sanders, 1966, and Hunkins, 1967), and analyze classroom questions (Davis and Tinsley, 1967, and Clegg et al. 1967a, 1967b).

Basically, the Taxonomy is made up of six categories of cognitive abilities:

(1) Knowledge - which involves the recall and remembering of information.

(2) Comprehension - which involves an understanding of that information rather than simple memory.

(3) Application - which involves putting the information to work.

(4) Analysis - which is a breakdown of the information into its integral parts and their logical organization.

(5) Synthesis - which is the joining together of thoughts to form new ideas.

(6) Evaluation - which is the making of a judgment about the value of something.
Pfeiffer and Davis (1965) reported a study in which Bloom's Taxonomy was used to determine the kind of objectives thought by teachers to be important. The investigators analyzed teacher-made tests. Results showed that test questions were mostly written at the first level (knowledge) of Bloom's Taxonomy. Social studies questions were written at an extremely low level. One interesting fact pointed out by the authors was that in civics classes all questions were at level one (knowledge) of the Taxonomy.

Davis and Hunkins (1966) reviewed the types of questions found in three different textbooks and decided that eighty-seven percent of the questions were the knowledge type. Hunkins (1967) followed up this study with one dealing with the construction of social studies textbook questions. He actually designed two sets of text-type materials. One contained forty-seven percent analysis and evaluation questions. The other had eighty-seven percent knowledge questions. Test results showed a significant advantage in social studies achievement favoring those students who were required to work at the analysis and evaluation levels.

Clegg et al. (1967a, 1967b) employed Bloom's Taxonomy in analyzing dialogue of elementary school children. Their study was patterned after the Davis-Tinsley study
(1967) which is discussed later. Clegg et al. actually trained student-teachers in the use of the Taxonomy as a means of raising the level of cognitive interaction. The cognitive level achieved by student-teachers in this study was significantly higher than those results reported by Davis and Tinsley. Clegg et al. also found that regular classroom teachers could be taught to identify the level of cognitive behavior that was currently being achieved in any classroom situation.

Working in the Manitowoc, Wisconsin School System, Sanders (1966) developed a technique for raising the level of oral classroom questioning. Sanders employed a modification of Bloom's Taxonomy which has seven categories: memory, translation, interpretation, application, analysis, synthesis, and evaluation (see Appendix A, p. 102). Sanders substituted "memory" for Bloom's heading "knowledge" as he felt it better identified the intellectual process involved. Sanders also divided Bloom's "comprehension" category into two of its subheadings, namely, "translation" and "interpretation." He felt that translation and interpretation offer two distinct kinds of thinking. In fact, Sanders attributed some operations to interpretation which Bloom placed under the analysis or the evaluation category. According to Sanders, the analysis process necessarily requires the individual
to employ the formal rules of logic. Most classroom activity does not involve such rigorous analysis. Instead, when the student is asked to compare two things, (and thus look at its parts) he does so only at a common-sense observational level. Sanders submitted that such behavior should be classified under the heading of interpretation rather than analysis.

To some degree, Sanders also differed with Bloom on the characteristics which make up the evaluation level. Bloom stated that the criteria upon which a judgment is based can be determined by the student or provided for him. Sanders held that the student must develop his own criteria for evaluation. According to Sanders, when the criteria are given to the student, only the interpretation level is reached.

Davis and Tinsley (1967) employed the Bloom-Sanders model to evaluate cognitive classroom interaction. They developed an instrument called a Teacher-Pupil Question Inventory (TPQI). It had nine categories to classify the types of questions that are asked by both teachers and pupils. The Bloom-Sanders Taxonomy made up the first seven categories. The remaining two categories, called "affectivity" and "procedure," covered non-cognitive situations.

The TPQI was used at the secondary level. An
observer was required to make a classroom visit of thirty minutes. At alternating five-minute intervals, he recorded the level of questions asked by the teacher or the student. Results of this experiment indicated that eighty-seven percent of teacher questions and ninety-six percent of student questions did not go above the third level (interpretation) of the Bloom-Sanders Taxonomy. The data also showed that memory was the major cognitive objective apparent in teachers' and pupils' verbal questions.

Conclusions

The Davis-Tinsley findings correspond with those of Bellack et al., Gallagher, and Clegg et al., reported earlier in this chapter. Together, they seem to point up the gap which exists between what educators claim should be the educational objectives and what are actually being achieved. The limited number of studies related to classroom cognitive questioning make it apparent that more investigations are needed in this area. Investigations should take the following forms:

(1) Replicate the present studies for verification purposes.

(2) Determine the optimal percentage of each category of questions in the teaching of different skills as well as the teaching of different
subjects.

(3) Determine if the Taxonomy is more useful at one level of education than it is at another.

(4) Determine if knowledge of the Taxonomy through pre-service and in-service training courses can effectively raise cognitive classroom behavior.

This present study concerned itself with items one and four, but was restricted to only that part of the classroom situation which may be classified as classroom questioning.
CHAPTER III
METHODS AND MATERIALS

As stated earlier, this study was designed to test the following hypotheses:

(1) There is no difference in the level of cognitive classroom behavior achieved by a group of student-teachers who have had instruction in the use of the Bloom-Sanders Taxonomy, and a group of student-teachers who have had no instruction in the use of the Taxonomy.

(2) There is no difference among investigators (R₁-R₃) in their rating of student-teachers, using the TPQI as a measuring device.

(3) There is no difference among co-operating teachers (R₄-R₉) in their rating of student teachers, using the TPQI as a measuring device.

(4) There is no difference in the level of cognitive classroom behavior achieved by student-teachers who have been instructed in the use of the Taxonomy.

(5) There is no increase over time in the level of cognitive classroom behavior of student-teachers
who have been instructed in the use of the Taxonomy.

(6) There is no difference between the rating scores of the investigators ($R_1-R_3$) and the rating scores of the co-operating teachers ($R_4-R_9$).

(7) There is no increase in rating agreement of co-operating teachers over time.

(8) There is no difference in the percentage of memory questions asked by a group of student-teachers who have had training in the use of the Taxonomy (experimental group) and a group of student-teachers who have had no training in the use of the Taxonomy (control group).

(9) There is no difference in the percentage of memory questions asked by a group of student-teachers who have had training in the use of the Taxonomy (experimental group) and a group of student-teachers previously reported in the literature who have had no training in the use of the Taxonomy (Davis and Tinsley, 1967).

Subjects

Two sample groups were used in this study. Twelve student-teachers (S) enrolled at the University of Massachusetts made up Group I. Six experienced teachers,
working in public school classrooms, who had student-teachers to supervise, made up Group II. These experienced teachers will be referred to hereafter as co-operating teachers.

None of the student-teachers had been introduced to the Bloom-Sanders Taxonomy (independent variable) during her formal training at the University. It was assumed, therefore, that the student-teachers had no knowledge of the Taxonomy. This assumption was substantiated about a week prior to the time that the student-teachers began their field work. On an examination, there was placed a question concerning the influence which Bloom's Taxonomy might have on the educational process. Over ninety percent of the students failed to respond to the question while the rest answered it in a manner which indicated their complete lack of knowledge of the subject. Furthermore, since no systematic biases were evident in the selection and placement of student-teachers, randomness was assumed.

Whether the co-operating teachers had prior knowledge of the Taxonomy is irrelevant due to the nature of the study and the research design that was employed. The question to be answered here was whether experienced teachers, using the TPQI, could agree on the level of cognitive behavior that was transpiring in the classroom.
This did not assume that experienced teachers had no previous knowledge of the Taxonomy. The co-operating teachers were also considered random samples from among the population of cooperating teachers used by the University. They became involved in this study by virtue of having a student-teacher (who was a subject) in their classroom.

Investigators

In an effort to diminish confounding conditions associated with experimenter bias, three raters \((R_1-R_3)\) were used to rate the taped lessons of the student-teachers in both the experimental and control groups. Together, these raters are referred to as the investigators. The investigators were doctoral candidates who held administrative positions in public school systems. Investigator number one \((R_1)\) instructed the student-teachers in the knowledge and procedures with which this study was concerned. He was charged with the responsibility of obtaining tapes of student-teachers' classroom lessons in social studies, having the co-operating teachers listen to and rate a random sequence from each tape, and arranging for the team of investigators to evaluate the same sequences later.

The investigators had previous experience in the use of the measuring instrument. Rater one was associated
with a pilot project (Clegg et al., 1967b) in which a similar instrument (TPQI) was used. Raters two and three were fellows, along with rater one, in the same research program at the University under which the pilot study was sponsored. Consequently, they were quite familiar with the Taxonomy and this method of investigation. Further training was given to them, however, before the study got under way. Practice in using the instrument was provided by employing classroom dialogue tapes obtained from the pilot study.

**Setting**

Two middle-class communities in western Massachusetts provided the necessary schools for this experiment. Six student-teachers were assigned to elementary schools in each community. One set of student-teachers was considered the experimental group, while the other set was the control group. In each community, there were student-teachers in two first grades, two second grades, and two third grades. The six co-operating teachers were faculty members from the school which had the experimental group of student-teachers. The co-operating teachers from the control group school were not involved in the experiment.
Experimental Procedures

During the first week of the student-teaching period, the investigator \( R_1 \) accompanied the supervisors of the student-teachers, assigned by the University, to the schools involved in the study. The supervisors introduced the investigator to the student-teachers and the co-operating teachers. The supervisors were helpful in creating a positive attitude toward this study by indicating their approval and interest in it. After the first week, the investigator worked independently of the supervisors and visited the schools each week on a day different from that of the supervisor.

Student Teachers

During the investigator's first visit, the student-teachers met with him for about an hour. On this occasion, the experimental group was given a summary sheet containing definitions of the terms found in the Taxonomy (see Appendix A, p. 102). The control group was given a summary sheet containing the ten categories which make up the Flanders-Amidon Model (see Appendix B, p. 104). By introducing the Flanders-Amidon model to the student-teachers of the control group, it was felt that they would consider themselves to be an experimental group. (In reality, they became a placebo). Since both groups
felt they were experimental groups, it was assumed that the presence of any Hawthorne Effect would have no greater influence upon the achievement of the experimental group over the control group. When these groups were compared later, it was further assumed that the influence of the Hawthorne Effect upon the experimental group was cancelled out by the Hawthorne Effect found in the placebo. Thus, while the magnitude of this effect was not assessed, its possible influence upon subjects within the study was treated.

Using, as a reference, the appropriate summary sheets, the investigator (R₁) explained to the student-teachers how the material was part of a study related to classroom analysis. Depending on the group, the investigator went on to explain either cognitive analysis or interaction analysis. He then requested each student-teacher to tape-record one social studies lesson each week. These taped lessons were reviewed weekly by the investigator, sometimes with the student-teachers (see Appendix C, p.106 for details of the weekly procedures used with both groups).

Also at the first meeting, a weekly schedule was presented to all twelve student-teachers. On some weeks, the investigator (R₁) met with the student-teachers as a group while at other times he met with them separately.
The schedule for the experimental group varied slightly from that of the control group since the independent variable (the Taxonomy) was used with one group and not with the other.

Group meetings for the experimental group centered around: (a) the problems that the student-teachers had in applying the Taxonomy, and (b) techniques which teachers might employ to raise and maintain the level of cognitive classroom behavior. Similar group meetings were held for the control group to discuss problems and techniques related to the Flanders-Amidon model. Meetings with individual student-teachers focused upon listening to the tapes which each had made. These sessions began with the student-teacher explaining the type of lesson she planned and the strategy she employed to obtain maximum effectiveness. She and the investigator (R₁) then analyzed the presentation from the standpoint of what was done well and how the lesson might possibly have been improved.

Co-operating Teachers

During the first week, the investigator (R₁) also met with the group of co-operating teachers. At this meeting, the teachers were given a weekly schedule and the same Taxonomy definitions that were given to the student-teachers of the experimental group. Most of
the time was spent in discussing the Taxonomy. The co-operating teachers were informed that, during the third, fifth, and seventh weeks, they would be asked to rate the six student-teachers, using the TPQI. (In the design, the co-operating teachers are identified as R₄-R₉). The material for the ratings of the student-teachers was obtained by playing to the co-operating teachers a randomly selected portion of each taped lesson for that week. A total of ten teacher questions and ten student responses in sequential order was used from each taped lesson. The same sequence was later rated by the three investigators (R₁-R₃). Tape recordings of lessons taught by the control group were rated at the same time intervals but by the investigators (R₁-R₃) only.

The investigator (R₁), on two occasions, met separately with each co-operating teacher. Depending upon what the teacher felt would be most valuable to her, the meetings proceeded in one of three ways: (1) the Taxonomy, in general, was discussed, (2) the student-teacher's application of the Taxonomy was discussed, or (3) part of the student-teacher's taped lesson was listened to and analyzed, either with or without the TPQI (see Appendix C, p.106).
Unit of Measurement

The criterion measure was the score on the Teacher-Pupil Question Inventory (TPQI) similar to the one employed by Clegg et al. (1967b). The TPQI is a single sheet which provides space for the rater to record the level of the teacher question and the level of the pupil response. (See Appendix D, p. 109).

Twenty scores (ten teacher questions and ten pupil responses) were recorded from each lesson. The sum of these scores was used as a composite achievement score of one student-teacher as determined by one rater.

Rules for Scoring the TPQI

In order that a uniform approach to classroom situations would be followed by all raters, the investigators (R₁-R₃) and co-operating teachers (R₄-R₉) were asked to observe the following rules:

1. Following Sanders' model the Analysis category of the Taxonomy was recorded only when the intellectual operation required the use of the formal rules of logic. Questions and responses which implied a "common sense" analysis were recorded at the Interpretation level.

2. Whenever the same question was asked of different pupils, it was recorded only once. This, in turn, meant that only one response for that
question was recorded. The recorded response was that which reached the highest cognitive level.

(3) "Yes" and "No" answers, if accepted by the teacher without pupil clarification, were recorded at the memory level.

Analysis of Data

The data required to test the hypotheses were obtained through use of the TPQI. Point values were assigned to the different cognitive levels associated with the TPQI. These values were:

- Memory ............. 1 point
- Translation ........ 2 points
- Interpretation ..... 3 points
- Application ........ 4 points
- Analysis ............ 5 points
- Synthesis ............ 6 points
- Evaluation ........... 7 points

As was mentioned earlier, each observer rated teacher questions and student responses and recorded twenty separate point scores for each student-teacher. The total of these twenty scores was used as the achievement score attributed to the student-teacher by one rater.

To test Hypotheses One and Two a four factor design with repeated measures was used. This included the
treatment variable (T) of whether the group did (T₁) or did not (T₂) have training in the use of the Taxonomy, the student-teacher variable (S), the rater variable (R), and the Trial variable (C). There were two treatments with six subjects assigned to each treatment. Each subject was rated three times by the same three raters. Analysis of variance was the statistical method employed. All factors, except subjects, were considered fixed.

Experimental Conditions associated with Hypotheses One and Two were:

Number of treatments (2)  T₁ ... T₂

Number of subjects (12)  S₁ ... S₆ (Experimental)

S'₁ ... S'₆ (Control)

Number of raters (3)  R₁ ... R₃ (Investigators)

Number of trials (3)  C₁ ... C₃ (Third, fifth, and seventh weeks)

The design associated with Hypotheses One and Two is shown in Figure 3. This same design was used in testing Hypotheses Eight and Nine.
For Hypotheses Three, Four, Five, Six, and Seven a three factor design with repeated measures was used to test results. The design included the six student-teachers of the experimental group (S), the three investigators and six co-operating teachers (R), and the three trial periods (C). Each subject was rated three times by each rater. Analysis of variance was the statistical method employed to test Hypotheses Three, Four, and Five. Another analysis of variance followed by orthogonal contrasts was employed to test Hypothesis Six. An F-test on variances was used to test Hypothesis Seven. Subjects and raters were considered as random factors, and trials as a fixed factor.
The experimental conditions associated with these hypotheses were:

Number of subjects (6) $S_1 \ldots S_6$ (Student-teachers)

Number of raters (9) $R_1 \ldots R_9$ (3 Investigators

$\begin{align*}
& R_1 \ldots R_3 \\
& 6 \text{ Co-operating teachers}
\end{align*}$

Number of trials (3) $C_1 \ldots C_3$

The Design associated with Hypotheses Three through Seven is given in Figure 4.

**FIGURE 4**

**DESIGN ASSOCIATED WITH HYPOTHESES THREE THROUGH SEVEN**
For Hypotheses Eight and Nine, the frequencies and percentages of questions asked at the different cognitive levels, were recorded in a table. The percentage of memory and above-memory questions asked by the experimental group was compared with the percentages found in (1) the control group, and (2) the Davis-Tinsley (1967) study. Chi-square statistical tests were used.

The frequency and percentage table is shown below:

**FIGURE 5**

FREQUENCY AND PERCENT OF QUESTIONS ASKED AT THE DIFFERENT COGNITIVE LEVELS

<table>
<thead>
<tr>
<th>Cognitive Levels</th>
<th>$S_1$ f %</th>
<th>$S_2$ f %</th>
<th>$S_3$ f %</th>
<th>$S_4$ f %</th>
<th>$S_5$ f %</th>
<th>$S_6$ f %</th>
<th>Total f %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Memory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Translation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Interpretation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Application</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Synthesis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Experimental Conditions associated with Hypothesis Eight were:

Control Group Percentages - $P_1$

Experimental Group Percentages - $P_2$
Experimental Conditions associated with Hypothesis Nine were:

- Davis-Tinsley - $P_1$
- Percentages

- Experimental Group - $P_2$
- Percentages

The chi-square table used to test both Hypothesis Eight and Hypothesis Nine was:

<table>
<thead>
<tr>
<th></th>
<th>$P_1$</th>
<th>$P_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above-Memory</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER IV

PRESENTATION AND ANALYSIS OF DATA

The analyses of data were performed in five steps. The first step consisted of an analysis of variance test which compared: (1) student-teacher mean scores of the experimental group with those of the control group (treatments); and (2) investigators' ($R_1$-$R_3$) mean scores. The second step consisted of an analysis of variance test to determine: (1) differences among co-operating teachers ($R_4$-$R_9$) as raters; (2) differences in cognitive classroom level of behavior achieved by student-teachers; and (3) effects of trials upon achievement. The third step was an analysis of variance on the nine raters (the three investigators and the six co-operating teachers). This was followed by an a priori contrast of investigators with co-operating teachers. The fourth step was the comparison of co-operating teachers' rating variances for each of the three trial periods. The fifth step was the compilation of frequency and percentage charts depicting the cognitive levels of student-teacher questions. The percentages were then used in a non-parametric test (chi-square) to determine if the experimental group asked more above-memory questions than either the control
group or the Davis-Tinsley (1967) group.

All data collected for this study are found in Appendices E and F on pages 112 and 113 respectively. Appendix E contains the composite scores necessary to test the first seven hypotheses. Appendix F, p. 113 contains the individual TPQI scores (questions and answers) of each student-teacher as determined by each rater. The frequency and percentage charts were made from the data appearing in Appendix F. The percentages associated with these charts were used to test Hypotheses Eight and Nine.

**Step One**

The hypotheses tested in this step were:

**Hypothesis One:** There is no difference in the level of cognitive classroom behavior achieved by a group of student-teachers who have had instruction in the use of the Bloom-Sanders Taxonomy, and a group of student-teachers who have had no instruction in the use of the Taxonomy.

**Hypothesis Two:** There is no difference among investigators (R₁-R₃) in their rating of student-teachers, using the TPQI as a measuring device.
The data used to test the hypotheses are shown in Table 1. This table is made up of the investigators' (R₁-R₃) ratings of both the experimental and control groups, as shown in Appendix E. Co-operating teachers' (R₄-R₉) rating scores are not included because they were not involved as supervisors of the control group. Table 2 contains results of this test.

Test results showed a significant difference (p<.01) between treatments. Therefore, Hypothesis One was rejected. It was concluded that teachers who have had instruction in the use of the Bloom-Sanders Taxonomy achieved a higher level of cognitive classroom behavior than student-teachers who have had no instruction in the Taxonomy.

The results also showed no difference among investigators. Therefore, Hypothesis Two was accepted. It was concluded that investigators, using the TPQI as a measuring device, agree in their rating of student-teachers.

Although only two hypotheses were stated here, the design permitted several other tests for multiple interaction to be made (as shown in Table 2). No differences were detected.
<table>
<thead>
<tr>
<th></th>
<th>T₁</th>
<th></th>
<th>T₂</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C₁</td>
<td>C₂</td>
<td>C₃</td>
</tr>
<tr>
<td>R₁</td>
<td>R₂</td>
<td>R₃</td>
<td>R₁</td>
</tr>
<tr>
<td>S₁</td>
<td>49 49 38</td>
<td>42 42 44</td>
<td>42 43 42</td>
</tr>
<tr>
<td>S₂</td>
<td>38 40 34</td>
<td>40 32 36</td>
<td>50 52 48</td>
</tr>
<tr>
<td>S₃</td>
<td>44 50 42</td>
<td>43 39 47</td>
<td>52 50 47</td>
</tr>
<tr>
<td>S₄</td>
<td>26 24 24</td>
<td>42 40 37</td>
<td>44 35 37</td>
</tr>
<tr>
<td>S₅</td>
<td>45 50 49</td>
<td>48 42 51</td>
<td>54 45 48</td>
</tr>
<tr>
<td>S₆</td>
<td>38 41 45</td>
<td>48 40 46</td>
<td>45 46 44</td>
</tr>
</tbody>
</table>

**TABLE 1**

INVESTIGATORS' (R₁-R₃) COMPOSITE SCORES FOR EXPERIMENTAL (T₁) AND CONTROL GROUP (T₂) OVER THREE TIME PERIODS (C₁-C₃)
<table>
<thead>
<tr>
<th>Sources of Variation</th>
<th>Degrees of Freedom</th>
<th>Sums of Squares</th>
<th>Mean Squares</th>
<th>Observed F Values</th>
<th>Critical F Values (.01)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Between</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatments (Fixed)</td>
<td>1</td>
<td>2223.15</td>
<td>2223.15</td>
<td>16.38*</td>
<td>10.0</td>
</tr>
<tr>
<td>Subjects within Treatments</td>
<td>10</td>
<td>1357.07</td>
<td>135.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Within</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investigators (Fixed)</td>
<td>2</td>
<td>13.50</td>
<td>6.75</td>
<td>.50</td>
<td>5.85</td>
</tr>
<tr>
<td>Investigators x Treatments</td>
<td>2</td>
<td>50.46</td>
<td>25.23</td>
<td>1.85</td>
<td>5.85</td>
</tr>
<tr>
<td>Investigators x Subjects within Treatments</td>
<td>20</td>
<td>272.48</td>
<td>13.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trials (Fixed)</td>
<td>2</td>
<td>147.17</td>
<td>73.58</td>
<td>1.28</td>
<td>5.85</td>
</tr>
<tr>
<td>Treatments x Trials</td>
<td>2</td>
<td>158.57</td>
<td>79.29</td>
<td>1.38</td>
<td>5.85</td>
</tr>
<tr>
<td>Trials x Subjects within Treatments</td>
<td>20</td>
<td>1149.37</td>
<td>57.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investigators x Trials</td>
<td>4</td>
<td>50.56</td>
<td>12.67</td>
<td>1.60</td>
<td>3.83</td>
</tr>
<tr>
<td>Treatments x Triads x Investigators x Trials</td>
<td>4</td>
<td>104.82</td>
<td>26.20</td>
<td>3.31</td>
<td>3.83</td>
</tr>
<tr>
<td>Treatments x Investigators x Subjects within Treatments</td>
<td>40</td>
<td>316.41</td>
<td>7.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td>107</td>
<td>5842.67</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant (p < .01)
Step Two

Three hypotheses were tested in this step. They were:

**Hypothesis Three**: There is no difference among co-operating teachers \((R_4-R_9)\) in their rating of student-teachers, using the TPQI as a measuring device.

**Hypothesis Four**: There is no difference in the level of cognitive classroom behavior achieved by student-teachers who have been instructed in the use of the **Taxonomy**.

**Hypothesis Five**: There is no increase over time in the level of cognitive classroom behavior of student-teachers who have been instructed in the use of the **Taxonomy**.

Table 3 contains the data which were used to test the above hypotheses. The table consists of the co-operating teachers' \((R_4-R_9)\) rating scores of the individual student-teacher. In Appendix E, p. 112, these scores appear under \(R_4\) through \(R_9\) of the experimental group. Investigators' scores are not included.

Test results related to the above hypotheses are shown in Table 4. This table shows a significant difference \((p. < .01)\) among co-operating teachers and among student-teachers. Therefore, Hypotheses Three and Four
TABLE 3
CO-OPERATING TEACHERS' \( (R_4-R_9) \) COMPOSITE
SCORES FOR EXPERIMENTAL GROUP \( (T_1) \)
OVER THREE TIME PERIODS \( (C_1-C_3) \)

<table>
<thead>
<tr>
<th></th>
<th>S_1</th>
<th>S_2</th>
<th>S_3</th>
<th>S_4</th>
<th>S_5</th>
<th>S_6</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R_4 )</td>
<td>C_1</td>
<td>C_2</td>
<td>C_3</td>
<td>C_1</td>
<td>C_2</td>
<td>C_3</td>
</tr>
<tr>
<td></td>
<td>53</td>
<td>40</td>
<td>41</td>
<td>37</td>
<td>35</td>
<td>36</td>
</tr>
<tr>
<td>( R_5 )</td>
<td>39</td>
<td>40</td>
<td>37</td>
<td>38</td>
<td>30</td>
<td>33</td>
</tr>
<tr>
<td>( R_6 )</td>
<td>52</td>
<td>53</td>
<td>38</td>
<td>36</td>
<td>30</td>
<td>36</td>
</tr>
<tr>
<td>( R_7 )</td>
<td>52</td>
<td>55</td>
<td>35</td>
<td>42</td>
<td>37</td>
<td>40</td>
</tr>
<tr>
<td>( R_8 )</td>
<td>47</td>
<td>43</td>
<td>38</td>
<td>38</td>
<td>35</td>
<td>48</td>
</tr>
<tr>
<td>( R_9 )</td>
<td>52</td>
<td>54</td>
<td>33</td>
<td>45</td>
<td>33</td>
<td>41</td>
</tr>
</tbody>
</table>

were rejected.

No difference could be detected among trials, and so Hypothesis Five was accepted. It was concluded, therefore, that (1) co-operating teachers, using the TPQI as a measuring device, differ in their rating of student-teachers; (2) there is a difference in the level of cognitive classroom behavior achieved by student-teachers who have been instructed in the use of the Taxonomy; and (3) there is no increase over time in the level of cognitive classroom behavior of student-teachers who have been instructed in the use of the Taxonomy.
TABLE 4

SUMMARY OF ANALYSIS OF VARIANCE TEST OF MEAN SCORES FOR CO-OPERATING TEACHERS, STUDENT-TEACHERS, AND TRIALS

<table>
<thead>
<tr>
<th>Sources of Variance</th>
<th>Degrees of Freedom</th>
<th>Sums of Squares</th>
<th>Mean Squares</th>
<th>Observed F Values</th>
<th>Critical F Values (.01)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-operating teachers (Random)</td>
<td>5</td>
<td>741.71</td>
<td>148.34</td>
<td>9.61*</td>
<td>3.86</td>
</tr>
<tr>
<td>Student-teachers (Random)</td>
<td>5</td>
<td>2869.93</td>
<td>573.98</td>
<td>37.19*</td>
<td>3.86</td>
</tr>
<tr>
<td>Co-operating teachers x Student-teachers</td>
<td>25</td>
<td>385.79</td>
<td>15.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trials (Fixed)</td>
<td>2</td>
<td>485.85</td>
<td>242.92</td>
<td>1.13</td>
<td>5.08</td>
</tr>
<tr>
<td>Co-operating teachers x Trials</td>
<td>10</td>
<td>96.37</td>
<td>9.64</td>
<td>.37</td>
<td>2.71</td>
</tr>
<tr>
<td>Student-teachers x Trials</td>
<td>10</td>
<td>2308.82</td>
<td>230.88</td>
<td>8.89*</td>
<td>2.71</td>
</tr>
<tr>
<td>Co-operating teachers x Student-teachers x Trials</td>
<td>50</td>
<td>1298.30</td>
<td>25.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>107</td>
<td>8186.77</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant (p. < .01)

After finding a difference among co-operating teachers as raters, it was desired to learn if this difference existed among many of the co-operating teachers or was
restricted to one or two. The Scheffé comparison test was subsequently employed. Results showed that only one co-operating teacher differed significantly from the others. Consequently, it can be stated that five of six co-operating teachers agreed in their rating of student-teachers.

A comparison test on the six student-teachers indicated that three had similar mean scores. Among the other three, one was significantly higher and two significantly lower.

The ANOVA table also shows a significant interaction of student-teachers x time, but no significant interaction of co-operating teachers x time. Co-operating teachers x student-teachers interaction could not be determined as there was no error term available to make the test.

The student-teachers x time interaction and the co-operating teachers x time interaction have been graphed below. Observation of the two graphs makes it clear why one interaction was significant while the other was not.
FIGURE 6
STUDENT-TEACHERS x TRIAL INTERACTION
(SIGNIFICANT)

FIGURE 7
CO-OPERATING TEACHERS x TRIAL INTERACTION
(NOT SIGNIFICANT)
Step Three

The hypothesis tested in this step was:

**Hypothesis Six**: There is no difference between
the rating scores of the investigators \((R_1 - R_3)\)
and the rating scores of the co-operating
teachers \((R_4 - R_9)\).

The data for this step are found in Table 5. The
table contains all of the scores for the experimental
group as shown in Appendix E, p. 112.

**TABLE 5**

RATERS' \((R_1 - R_9)\) COMPOSITE SCORES FOR
EXPERIMENTAL GROUP \((T_1)\) OVER
THREE TIME PERIODS \((C_1 - C_3)\)

<table>
<thead>
<tr>
<th></th>
<th>(S_1)</th>
<th>(S_2)</th>
<th>(S_3)</th>
<th>(S_4)</th>
<th>(S_5)</th>
<th>(S_6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(R_1)</td>
<td>C_1C_2C_3</td>
<td>C_1C_2C_3</td>
<td>C_1C_2C_3</td>
<td>C_1C_2C_3</td>
<td>C_1C_2C_3</td>
<td>C_1C_2C_3</td>
</tr>
<tr>
<td></td>
<td>49 42 42</td>
<td>38 40 50</td>
<td>44 43 52</td>
<td>26 42 44</td>
<td>45 48 54</td>
<td>38 48 45</td>
</tr>
<tr>
<td>(R_2)</td>
<td>49 42 43</td>
<td>40 32 52</td>
<td>50 39 50</td>
<td>24 40 35</td>
<td>50 42 45</td>
<td>41 40 46</td>
</tr>
<tr>
<td>(R_3)</td>
<td>38 44 42</td>
<td>34 36 48</td>
<td>42 47 47</td>
<td>24 37 37</td>
<td>49 51 48</td>
<td>45 46 44</td>
</tr>
<tr>
<td>(R_4)</td>
<td>53 40 41</td>
<td>37 35 36</td>
<td>45 48 50</td>
<td>26 39 41</td>
<td>43 47 41</td>
<td>32 52 34</td>
</tr>
<tr>
<td>(R_5)</td>
<td>39 40 37</td>
<td>38 30 33</td>
<td>43 54 52</td>
<td>26 40 42</td>
<td>49 41 41</td>
<td>33 60 40</td>
</tr>
<tr>
<td>(R_6)</td>
<td>52 53 38</td>
<td>36 30 36</td>
<td>53 60 53</td>
<td>29 41 40</td>
<td>46 46 41</td>
<td>44 50 42</td>
</tr>
<tr>
<td>(R_7)</td>
<td>52 55 35</td>
<td>42 37 40</td>
<td>63 50 56</td>
<td>29 46 43</td>
<td>39 64 57</td>
<td>47 62 44</td>
</tr>
<tr>
<td>(R_8)</td>
<td>47 43 38</td>
<td>38 35 48</td>
<td>45 47 55</td>
<td>32 48 46</td>
<td>43 48 51</td>
<td>29 61 41</td>
</tr>
<tr>
<td>(R_9)</td>
<td>52 54 33</td>
<td>45 33 41</td>
<td>54 45 57</td>
<td>27 41 41</td>
<td>46 52 60</td>
<td>41 54 47</td>
</tr>
</tbody>
</table>
The statistical design made it possible to test differences among raters \((R_1-R_9)\). The design also permitted a re-test of Hypotheses Four and Five which were related to student-teacher effects and trial effects. Table 6 contains the results.

**TABLE 6**

**SUMMARY OF ANALYSIS OF VARIANCE TEST OF MEAN SCORES FOR RATERS \((R_1-R_9)\)**

<table>
<thead>
<tr>
<th>Sources of Variation</th>
<th>Degrees of Freedom</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>Observed F Values</th>
<th>Critical F Values (.01)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Raters</td>
<td>8</td>
<td>1272.00</td>
<td>159.00</td>
<td>18.70*</td>
<td>2.99</td>
</tr>
<tr>
<td>(Random)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student-teachers</td>
<td>5</td>
<td>3567.77</td>
<td>713.55</td>
<td>83.90*</td>
<td>3.51</td>
</tr>
<tr>
<td>(Random)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raters x Student-teachers</td>
<td>40</td>
<td>340.90</td>
<td>8.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within Trials</td>
<td>2</td>
<td>495.44</td>
<td>247.72</td>
<td>0.98</td>
<td>4.93</td>
</tr>
<tr>
<td>(Fixed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raters x Trials</td>
<td>16</td>
<td>46.89</td>
<td>2.93</td>
<td>.10</td>
<td>2.31</td>
</tr>
<tr>
<td>Student-teachers x</td>
<td>10</td>
<td>2521.90</td>
<td>252.19</td>
<td>8.28*</td>
<td>2.59</td>
</tr>
<tr>
<td>Trials</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raters x Student-teachers x Trials</td>
<td>80</td>
<td>2347.10</td>
<td>29.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>161</td>
<td>10592.00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant \((p < .01)\).*
The findings showed a significant difference (p < 0.01) among raters (R₁-R₉). There was also a significant difference (p < 0.01) among student-teachers and no significant difference among trials. (These last two findings correspond with what has been reported earlier in Step Two).

Since Hypothesis Six implies a test between groups (R₁-R₂ vs R₄-R₉) and since a significant difference was found among individual raters, there was, at this point, insufficient information to draw a conclusion. Therefore, an a priori comparison was performed next. The results showed a significant difference between groups. Consequently, Hypothesis Six was rejected. It was concluded that there is a difference between the rating scores of investigators and co-operating teachers. Subsequently, a second contrast (Scheffé) was performed using five of the co-operating teachers. The one teacher who earlier had been found to be different was not included in this contrast. The new results indicated no difference between the three investigators and the five co-operating teachers.

**Step Four**

The hypothesis tested in this step was:

**Hypothesis Seven**: There is no increase in rating agreement of co-operating teachers over time.
To test this hypothesis, co-operating teacher rater \((R_4 - R_9)\) variances for the three trail periods were compared. The data necessary to obtain rater variances are contained in Table 3.

Test results for this hypothesis are given in Table 7.

**TABLE 7**

**CO-OPERATING TEACHERS VARIANCES FOR THE THREE TRIAL PERIODS**

<table>
<thead>
<tr>
<th>Trial</th>
<th>Observed (S^2)</th>
<th>Critical (F(.05))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.64</td>
<td>1.13</td>
</tr>
<tr>
<td>2</td>
<td>9.69</td>
<td>1.13</td>
</tr>
<tr>
<td>3</td>
<td>8.56</td>
<td></td>
</tr>
</tbody>
</table>

The findings indicate no significant difference among variances. Therefore, Hypothesis Seven was accepted. It was concluded that there is no increase in rating agreement of co-operating teachers over time.
Step Five

The hypotheses tested in this step were:

Hypothesis Eight: There is no difference in the percentage of memory questions asked by a group of student-teachers who have had training in the use of the Taxonomy (experimental group) and a group of student-teachers who have had no training in the use of the Taxonomy (control group).

Hypothesis Nine: There is no difference in the percentage of memory questions asked by a group of student-teachers who have had training in the use of the Taxonomy (experimental group) and a group of student-teachers, previously reported in the literature, who had no training in the use of the Taxonomy (Davis and Tinsley, 1967).

First, frequencies and percentages of the cognitive levels of student-teacher questions, were compiled from the data contained in Appendix F, p. 113. These are shown in Table 8. Since only the investigators ($R_1-R_3$) had evaluated both the experimental and control groups of student-teachers, the investigators' ratings were used. Frequencies and percentages were obtained by tabulating the number of questions asked by the student-
TABLE 8
FREQUENCY AND PERCENTAGE OF STUDENT-TEACHERS QUESTIONS ARRANGED BY COGNITIVE LEVELS

Control Group (R₁-R₃)

<table>
<thead>
<tr>
<th>Cognitive Levels</th>
<th>S₁</th>
<th>S₂</th>
<th>S₃</th>
<th>S₄</th>
<th>S₅</th>
<th>S₆</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Memory</td>
<td>28</td>
<td>31</td>
<td>33</td>
<td>37</td>
<td>43</td>
<td>48</td>
<td>16 18</td>
</tr>
<tr>
<td>2 Translation</td>
<td>46</td>
<td>51</td>
<td>46</td>
<td>51</td>
<td>32</td>
<td>36</td>
<td>58 64</td>
</tr>
<tr>
<td>3 Interpretation</td>
<td>13</td>
<td>14</td>
<td>11</td>
<td>12</td>
<td>15</td>
<td>17</td>
<td>16 18</td>
</tr>
<tr>
<td>4 Application</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0   0</td>
</tr>
<tr>
<td>5 Analysis</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0   0</td>
</tr>
<tr>
<td>6 Synthesis</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0   0</td>
</tr>
<tr>
<td>7 Evaluation</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0   0</td>
</tr>
</tbody>
</table>

Experimental Group (R₁-R₃)

<table>
<thead>
<tr>
<th>Cognitive Levels</th>
<th>S₁</th>
<th>S₂</th>
<th>S₃</th>
<th>S₄</th>
<th>S₅</th>
<th>S₆</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Memory</td>
<td>9</td>
<td>10</td>
<td>22</td>
<td>24</td>
<td>15</td>
<td>17</td>
<td>36 40</td>
</tr>
<tr>
<td>2 Translation</td>
<td>43</td>
<td>48</td>
<td>45</td>
<td>50</td>
<td>29</td>
<td>32</td>
<td>37 41</td>
</tr>
<tr>
<td>3 Interpretation</td>
<td>36</td>
<td>40</td>
<td>23</td>
<td>26</td>
<td>42</td>
<td>47</td>
<td>14 16</td>
</tr>
<tr>
<td>4 Application</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1   1</td>
</tr>
<tr>
<td>5 Analysis</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0   0</td>
</tr>
<tr>
<td>6 Synthesis</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0   0</td>
</tr>
<tr>
<td>7 Evaluation</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>2   2</td>
</tr>
</tbody>
</table>
teachers at each cognitive level.

Percentages of memory questions were then used to contrast, by chi-square, the experimental group with: (1) the control group; and (2) the Davis-Tinsley group. Figures for the Davis-Tinsley group were already available. They are 61.3 percent memory and 38.7 percent above-memory.

Test results related to Hypotheses Eight and Nine are given in Table 9.

TABLE 9

<table>
<thead>
<tr>
<th></th>
<th>Memory Level</th>
<th>Above-Memory</th>
<th>Chi-Square</th>
<th>Critical Value (.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control Group</td>
<td>31.5</td>
<td>68.5</td>
<td>5.98*</td>
<td>3.84</td>
</tr>
<tr>
<td>Experimental Group</td>
<td>16.7</td>
<td>83.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Memory Level</th>
<th>Above-Memory</th>
<th>Chi-Square</th>
<th>Critical Value (.001)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Davis-Tinsley</td>
<td>61.3</td>
<td>38.7</td>
<td>41.8**</td>
<td>10.827</td>
</tr>
<tr>
<td>Experimental Group</td>
<td>16.7</td>
<td>83.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant (p<.05)
**Significant (p<.001)
The results showed a significant difference \((p<0.05)\) between the control group and the experimental group and a significant difference \((p<0.001)\) between the Davis-Tinsley group and the experimental group. Therefore, Hypotheses Eight and Nine were rejected. It was concluded that student-teachers who have training in the use of the Taxonomy as a teaching device ask more above-memory questions than student-teachers who have no training in the use of the Taxonomy.

Closer observation of the individual TPQI's (Appendix F) revealed an interesting point. When the levels of pupil responses and student-teacher questions were combined to form another frequency and percentage chart (similar to Table 8) the overall cognitive level was lowered. Frequencies and percentages relative to this are presented in Table 10 so that a comparison may be made. Table 11 was devised so that total frequencies and percentages related to Tables 8 and 10 might easily be compared. Compare, in Table 11 for example, level one of the experimental groups. For student-teacher questions only, there is a percentage of 16.7 for the memory category. For student-teacher questions and pupil responses combined, there is a percentage of 25.0 for the same category. Such a situation indicates that a pupil's response is often at a lower cognitive level.
TABLE 10
FREQUENCY AND PERCENTAGE OF COMBINED STUDENT-TEACHER QUESTIONS AND PUPIL RESPONSES ARRANGED BY COGNITIVE LEVELS

Control Group (R₁⁻R₃)

<table>
<thead>
<tr>
<th>Cognitive Levels</th>
<th>S₁</th>
<th>S₂</th>
<th>S₃</th>
<th>S₄</th>
<th>S₅</th>
<th>S₆</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>f%</td>
<td>f%</td>
<td>f%</td>
<td>f%</td>
<td>f%</td>
<td>f%</td>
<td>f%</td>
</tr>
<tr>
<td>1 Memory</td>
<td>90</td>
<td>50</td>
<td>81</td>
<td>45</td>
<td>79</td>
<td>44</td>
<td>10458</td>
</tr>
<tr>
<td>2 Translation</td>
<td>72</td>
<td>40</td>
<td>86</td>
<td>48</td>
<td>67</td>
<td>37</td>
<td>11662</td>
</tr>
<tr>
<td>3 Interpretation</td>
<td>15</td>
<td>8</td>
<td>13</td>
<td>7</td>
<td>31</td>
<td>17</td>
<td>2716</td>
</tr>
<tr>
<td>4 Application</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5 Analysis</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6 Synthesis</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7 Evaluation</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Experimental Group (R₁⁻R₃)

<table>
<thead>
<tr>
<th>Cognitive Levels</th>
<th>S₁</th>
<th>S₂</th>
<th>S₃</th>
<th>S₄</th>
<th>S₅</th>
<th>S₆</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>f%</td>
<td>f%</td>
<td>f%</td>
<td>f%</td>
<td>f%</td>
<td>f%</td>
<td>f%</td>
</tr>
<tr>
<td>1 Memory</td>
<td>35</td>
<td>21</td>
<td>59</td>
<td>27</td>
<td>43</td>
<td>24</td>
<td>89149</td>
</tr>
<tr>
<td>2 Translation</td>
<td>81</td>
<td>45</td>
<td>96</td>
<td>53</td>
<td>61</td>
<td>34</td>
<td>10257</td>
</tr>
<tr>
<td>3 Interpretation</td>
<td>59</td>
<td>33</td>
<td>29</td>
<td>16</td>
<td>70</td>
<td>38</td>
<td>5128</td>
</tr>
<tr>
<td>4 Application</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5 Analysis</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6 Synthesis</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7 Evaluation</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>16</td>
</tr>
</tbody>
</table>
TABLE 11
COMPARISON OF TOTALS OF FREQUENCY AND PERCENTAGE
TEACHER QUESTIONS AND COMBINED TEACHER QUESTIONS
AND PUPIL RESPONSES ARRANGED BY COGNITIVE
LEVELS (FROM TABLES 8 AND 10)

<table>
<thead>
<tr>
<th>Cognitive Level</th>
<th>Control</th>
<th>%</th>
<th>Experimental</th>
<th>Control</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Memory</td>
<td>170</td>
<td>32</td>
<td>90</td>
<td>271</td>
<td>25</td>
</tr>
<tr>
<td>2 Translation</td>
<td>282</td>
<td>52</td>
<td>234</td>
<td>466</td>
<td>43</td>
</tr>
<tr>
<td>3 Interpretation</td>
<td>83</td>
<td>15</td>
<td>204</td>
<td>319</td>
<td>30</td>
</tr>
<tr>
<td>4 Application</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>5 Analysis</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6 Synthesis</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7 Evaluation</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>16</td>
<td>2</td>
</tr>
</tbody>
</table>

Cognitive Level | Combined Teacher Questions and Pupil Responses (from Table 10) |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Memory</td>
<td>468 43</td>
</tr>
<tr>
<td>2 Translation</td>
<td>491 46</td>
</tr>
<tr>
<td>3 Interpretation</td>
<td>115 11</td>
</tr>
<tr>
<td>4 Application</td>
<td>6 1</td>
</tr>
<tr>
<td>5 Analysis</td>
<td>0 0</td>
</tr>
<tr>
<td>6 Synthesis</td>
<td>0 0</td>
</tr>
<tr>
<td>7 Evaluation</td>
<td>0 0</td>
</tr>
</tbody>
</table>
than the teacher's question. Further analysis of the tables revealed that seventeen percent of the time when above-memory questions were asked of the experimental group, the answers were at the memory level. Within the control group, the percentage was even higher as twenty-four percent of above-memory questions were answered at the memory level.

It should be noted in Table 11 that at Level 2 (Translation) of the experimental group, the percentages are almost identical. This does not mean, however, that the level of student responses was always consistent with the level of teachers questions within the translation category. A broader comparison of the tables shows that within the first three categories (memory, translation, and interpretation) there was a downward trend of about eight percent from Table 8 to Table 10. Consequently, it may appear that there was no change at Level 2 when actually this level lost about eight percent to the level below but gained about eight percent from the level above. Such a condition can be verified by referring to the data in Appendix F.

The above findings may be partially explained by one of the ground rules by which the lessons were analyzed. The raters had agreed beforehand that when a "Yes" or "No" answer went unchallenged by the teacher,
a rating of "1" would be given to that response. This condition would exist regardless of the possibility that higher level thinking had occurred. Such a procedure probably had a significant effect upon the percentage differences found in the memory category as opposed to the above-memory categories. Nevertheless, the differences between Table 8 and Table 10 are great enough to indicate that students are not always operating at as high a cognitive level as the level of teacher questions would appear to indicate. These findings are in contrast to those previously reported by Taba (1967) and Clegg (1967b) and warrant further investigation.
Results of this experiment and their implications are considered in this chapter under six headings. The first three main hypotheses are discussed under separate headings, followed in order by a discussion of secondary hypotheses, general conclusions, and recommendations.

Discussion of Hypothesis One

Hypothesis One was tested in order to obtain an answer to the following question: can instruction in the use of the Taxonomy aid student-teachers in asking higher cognitive classroom questions and in eliciting higher cognitive pupil responses? Test results provide an affirmative answer to this question. When two groups of student-teachers were compared, it was found that those who had been trained in the use of the Taxonomy achieved higher cognitive classroom behaviors. This conclusion is supported by the findings related to Hypotheses Eight and Nine. In these hypotheses, the percentages of memory questions asked by different groups of student-teachers were compared. It was found that those who had instruction in the use of the Taxonomy asked more above-
memory questions. Thus, it was concluded that student-teachers who have training in the use of the Taxonomy achieve higher cognitive classroom behaviors than the student-teachers who have no training in the use of the Taxonomy.

These results correspond with those reported by Hunkins (1967). Both this study and the Hunkins study found higher cognitive achievement in social studies as a result of employing the Taxonomy. It seems desirable, therefore, that the classroom teacher have a knowledge of the Taxonomy and be instructed in its use as a tool for raising the cognitive level of classroom behavior.

A closer look at the frequency and percentage tables associated with Hypotheses Eight and Nine indicated, however, that seldom did the questions and responses go above the interpretation level. It was also apparent that pupils often answered questions at a lower cognitive level than at which the question was asked. This finding seems to support the inference drawn from Klebaner (1964) that pupils should be made aware of the type of answer that is expected of them so that their responses will be at the desired cognitive level.

Overall, the findings presented here are quite similar to those of Gallagher (1965) who found that most questions fell into the cognitive-memory category, and
Davis and Tinsley (1967), who reported that eighty-seven percent of student-teacher questions did not go above the interpretation level.

It must be concluded, therefore, that this study was successful in raising cognitive classroom behavior, but only at the lower intellectual levels. The findings reported here coupled with those of Gallagher, and Davis and Tinsley support the charge that student-teachers ask very few challenging or thought-provoking questions, and that the art of questioning is still in a primitive stage. This seems to imply that teachers need to be provided with (1) a means of clearly identifying the levels of intellectual thought; and (2) a method of asking questions which stimulates higher cognitive processes.

The findings suggest that instruction in the use of the Taxonomy as a tool for clearly identifying the different cognitive levels will be helpful to both student-teachers and experienced teachers alike. Thus, it is recommended that the Taxonomy be included in the curriculum for pre-service education of teachers and also within the in-service program of continuing teacher education. In addition, introducing teachers to the Taxonomy and the technique of using tapes and the TPQI as an instrument for analysis will have a three-fold advantage: (1) it will give teachers the means to
evaluate their course objectives to determine whether higher level objectives have been achieved; (2) it will provide teachers with a method of including objectives in daily lesson plans that stimulate higher levels of learning; and (3) it will make teachers more alert to various impromptu classroom situations and conditions that provide opportunities for developing higher levels of thinking.

Two methods in which the Taxonomy may be meaningfully introduced to student-teachers seem possible. One is to combine the teaching of the Taxonomy with the questioning technique related to Taba's (1967) schema of cognitive tasks. Such an approach would allow the student-teacher to clearly perceive the desired cognitive level and provide him with an effective teaching-learning objective to bring about the intellectual behavior necessary to reach that level.

A second approach is to use the Taxonomy with micro-teaching. The micro-lesson could be recorded on audio-tape or, if materials and funds are available, on video-tape. Lesson objectives could be written at various levels of the Taxonomy. Then, one cognitive objective from the lesson plan could be focused upon in the analysis of the micro-lesson and a decision made on how effectively that objective was achieved. Whenever
teacher-pupil questioning were used as the teaching method, the TPQI could be employed as an evaluating instrument.

Although the remarks above were directed toward improvement of student-teacher education, similar approaches could be used with experienced teachers. Workshop courses could easily be developed where teachers would plan lessons for their own classes, tape record the lessons and evaluate them afterwards using the **Taxonomy**. In some cases, these workshops could take the form of micro-teaching using portable video-tape facilities.

**Discussion of Hypothesis Two**

The second main question, which was tested under Hypothesis Two, was: can investigators, when using the TPQI, agree upon the cognitive classroom behavior that is being observed? Since test results showed no significant difference among rating scores for \((R_1-R_2)\), it was concluded that investigators can agree in their observations of cognitive classroom behavior. This is a most important finding as it substantiates the basic assumption of this study. It had been assumed that investigators' ratings can be relied upon for the purpose of making comparisons, such as treatment methods and achievement of student-teachers.
The low F value associated with the test of investigators' \((R_1 - R_2)\) ratings suggests high reliability within the TPQI. This fact is important as it adds further confidence in other test findings. It also adds credence to the frequency and percentage charts which are based entirely on the investigators' ratings scores.

That the TPQI apparently has high reliability coupled with the fact that it is easy to use suggests that it might be attractive to supervisors of both student-teachers and experienced teachers. In either case, the supervisor could employ this simple device to determine the cognitive behavior stimulated by the classroom teacher.

The Taxonomy itself might be incorporated into one of the standard measuring devices, such as the OScAR. At present, the only attempt within the OScAR to determine cognitive behavior is related to a scoring key which measures "problem-structuring of teacher statements." As a multi-dimensional record, the OScAR has three subclassifications: (1) emotional climate; (2) verbal emphasis; and (3) social organization. It would seem that a fourth classification, cognitive behavior, would fit into this scheme. A cognitive behavior section is most desirable and, as part of the OScAR, would add much in obtaining a true and complete picture of a teacher's
Another device with which the Taxonomy could be used is Flanders' interaction analysis. The Taxonomy might be used jointly with this technique or, as Clegg (1967b) has suggested, incorporated into the matrix itself. Either process might yield important information relating to the effects of different types of teacher-pupil interaction on cognitive development. Correlation of interaction analysis scores with TPQI scores might provide an answer to the question: does teaching style make a difference? It would seem that investigations in this area are most desirable.

Recently, Amidon (1968) reported that he and his associates have made attempts to add a cognitive dimension to the Flander's matrix. Amidon used the categories of Gallagher's model as sub-headings under teacher statements and pupil statements found in the original 10 x 10 matrix. The matrix contains over twenty categories. At this time, Amidon was unable to report any outstanding results with the revised instrument. A major problem was related to recording procedures in which a tally mark is to be placed in the matrix every three seconds. The evaluator often found it difficult to keep pace. Not only must he observe who is speaking but now must also decide upon the cognitive level of the statement.
Discussion of Hypothesis Three

The third main question of the study was: Can co-operating teachers, as raters, agree upon the cognitive classroom behavior that is being observed? Hypothesis Three, relative to this question, was rejected. Thus, it was concluded that the co-operating teachers \((R_4-R_9)\) could not agree in their rating scores. Although Hypothesis Three was rejected, it was encouraging to note that five of six raters were able to agree in their evaluations. The test of Hypothesis Three was a replication of the Clegg study (1967b) where complete agreement of co-operating teachers had been achieved.

Hypothesis Three closely parallels Hypothesis Six and thus, it seems logical that the findings related to Hypothesis Six should be briefly discussed at this point. Hypothesis Six states that there was no difference between the mean rating scores of co-operating teachers and the mean rating scores of investigators. Test results showed that a significant difference did exist between the groups. It was subsequently found by a Sheffé contrast, however, that only one co-operating teacher (the same person identified as being different in Hypothesis Three) differed significantly from the other eight raters. Thus, it can be stated that eight of nine raters agreed in their evaluation of student-
teachers.

It is felt, therefore, that a conclusive answer to the above question has not yet been obtained. Prospects seem likely, however, that co-operating teachers can agree upon observed cognitive classroom behaviors when they employ the Taxonomy as a rating device.

If Hypothesis Three can be accepted, such a finding could be very valuable for the improvement of instruction. It would provide the co-operating teacher with a means to quickly and clearly identify the cognitive behavior that is being developed by the student-teacher. It would mean that the University supervisor, the co-operating teacher, and the student-teacher could discuss a phase of the educational process in terms that are understandable to all and free of possible subjective bias. In other words, the Taxonomy could serve as a common language for evaluation of instruction, as Clegg (1967b) has already suggested.

**Discussion of Secondary Hypotheses**

Hypothesis Four concerning equal achievement of student-teachers, was rejected. This finding can be attributed to individual differences. The task that is implied is how to improve the effectiveness of each teacher involved, especially those whose level of
questioning appears consistently low.

Hypothesis Five, concerning the lack of increase in student-teacher achievement over time, was accepted. This deserves some consideration. The findings seem to indicate that all that was achieved in this experiment was accomplished after three weeks, and that what happened thereafter was only repetition. This suggests that greater success might be forthcoming if the experiment were conducted in three phases, in a manner that would allow greater concentration at a single cognitive level. Each phase could correspond to the three time periods established for the rating of the tapes (at the end of the third, fifth, and seventh weeks). Phase one might be carried on just as before, with emphasis placed on encouraging student-teachers to ask questions which go beyond the memory level. Phase two might concentrate on encouraging student-teachers to ask application questions, each of which usually has one best answer. Phase three would concentrate on synthesis and evaluation questions which often lead to unique answers. This approach would emphasize convergent thinking during phase two and divergent thinking during phase three. Since the Taxonomy is thought of as hierarchical, no phase of the cognitive process would be overlooked. This suggested approach differs from the one used in the present experiment where
the Taxonomy was always looked at as a whole and no level was given special consideration.

Hypothesis Six, which stated that there was no difference between the mean scores of investigators and those of co-operating teachers, was rejected. Implications related to this hypothesis were discussed earlier under Hypothesis Three. It should be recalled that agreement of rating scores was found among eight of the nine raters. No definite conclusions were made as the findings here conflicted with those reported by Clegg (1967b) who found complete agreement among co-operating teachers' rating scores. If, in the future, it can be shown with increased n's that experienced teachers and supervisors can agree on the level of observed cognitive behavior, the Taxonomy can be used as a common language in discussing cognitive levels of achievement.

Hypothesis Seven, relative to lack of increase in agreement of co-operating teachers' rating scores over time, was accepted. The observation can be made that the rating variance for trial three was lower (but not significantly) than the other two variances. To imply that a linear trend was developing, however, would be treading on dangerous ground indeed. Therefore, it must be concluded that the findings are too limited to suggest anything about the failure of co-operating
teachers' rating scores to increase in agreement over time.

Hypotheses Eight and Nine, concerned with comparisons of the percentages of memory and above-memory questions asked by the experimental group with percentages asked by two other groups who did not employ the Taxonomy, were rejected. It was found that the experimental group asked a larger percentage of above-memory questions. These results and their implications were discussed earlier as they were intrinsically related to Hypothesis One. The conclusion was that although the experimental group asked more above-memory questions, seldom did the level of questioning go above the interpretation category. Thus, it was felt that the findings were similar to those of Gallagher, and Davis and Tinsley.

When a frequency and percentage table, using the nine rater scores, was compared with a similar table from the Clegg study (1967b), the results for the lower categories were relatively the same. The difference at the memory level was only five percent favoring the Clegg group. This difference does not approach significance. There is a marked difference of percentages at the higher levels, however, where Clegg reported a total of thirty-five percent for analysis, synthesis and evaluation questions. The present study reported only
one and one-half percent for analysis, synthesis and evaluation. A partial explanation for this difference is to be found in the model that was employed. It should be recalled that Clegg used Bloom's Taxonomy which has six categories, whereas, this study used Sander's modification of Bloom's Taxonomy which has seven categories several of which are defined more rigorously than the original Bloom model. Most important is the fact that Sanders: (1) considered most analysis questions at a lower level (interpretation); and (2) held that evaluation questions occur, only, when the student generates his own values. In contrast, Bloom held that evaluation may also occur when the student uses an externally derived set of values as a criterion.

The results of this study give rise to several other questions related to cognitive development. It will be noted that this study concerned itself with children in grades one through three. The findings compared favorably with those of Davis and Tinsley who worked with secondary school students. Although the findings are tentative, they seem to indicate that children in the early grades can operate at the same cognitive levels (although intellectual sophistication may differ) as children in the secondary grades. Clegg (1967b) arrived at a similar conclusion. It seems important, therefore,
to obtain an answer to the following question: Is there a difference in the levels of cognitive behavior among children in elementary schools compared to children in secondary schools?

**General Conclusions**

Several general conclusions can be stated as a result of this study. First, instruction in the use of the **Taxonomy** as a tool for clearly identifying the different levels of intellectual behavior will be helpful to both student-teachers and experienced teachers alike. Second, since the TPQI is simple to use and appears to have high reliability, it may be a useful device for supervisors of teachers. Third, close agreement by raters on the cognitive behavior occurring in the classroom, suggests that the **Taxonomy** can serve as a common educational language for improvement of instruction.

Finally, it must be conceded that social study goals which call for higher levels of thinking are not being achieved. The findings reported here are similar to those of Gallagher, and Davis and Tinsley. Together, they challenge the time-honored practice of asking questions as a method of teaching. A corollary to the above is that teachers are not being given the proper training to successfully implement the higher levels of
thinking in the classrooms.

**Recommendations**

The findings reported herein appear to suggest a promising approach toward the improvement of teacher education and pupil achievement. It appears desirable to include study of the *Taxonomy* in the teacher-training curriculum as a tool for raising the level of classroom learning. The following recommendations are suggested for introducing the *Taxonomy* to teachers in a meaningful way: (1) combine the teaching of the *Taxonomy* with Taba's (1967) schema of cognitive tasks; (2) use the *Taxonomy* with micro-teaching.

The success of the TPQI as a measurement of cognitive behavior suggests that the *Taxonomy* could be used with other measuring instruments. The *Taxonomy* might be used to add a cognitive dimension to the OScAR, or it might be employed with Flanders' interaction analysis to determine the effect of teaching style on cognitive achievement.

In view of the small sample size, this study should be replicated on a larger scale to verify the findings and conclusions presented here. A suggested modification might be to teach the *Taxonomy* in three phases. Phase one would concentrate on encouraging teachers to ask above-memory questions. Phase two would emphasize
convergent thinking (application questions). Phase three would emphasize divergent thinking (analysis, synthesis, and evaluation).

Several other questions might also be tested in subsequent studies. For example: (1) Will the cognitive response level increase if the pupil is taught to become aware of the level of his expected response? (2) Does eliciting longer answers from the pupil tend to raise the cognitive level of the response? (3) Is there a correlation between the type of teacher (according to the Flanders model) and the cognitive level attained in the classroom (as measured by the Taxonomy)? (4) Do students operate at the same cognitive level in one subject as they do in another?
CHAPTER VI

SUMMARY

Statement of Problem

One of the major problems in education is that high-level goals, such as increasing the level of intellectual behavior, are often stated as educational objectives but are seldom achieved by classroom teachers. Available evidence suggests that the majority of classroom questions are limited to memory-type learning (Gallagher, 1965), and that questions at the memory and comprehension levels make up over eighty-five percent of all classroom questions (Davis and Tinsley, 1967). Since classroom questioning is an integral part of the overall cognitive classroom behavior, it is essential that methods be found that will enable teachers to raise the cognitive level of their questions.

Related Research

Since 1914, several approaches have been taken toward increasing teacher effectiveness. This writer has chosen to classify them under four headings: (1) early studies, which evaluated student statements or teacher statements to determine the distribution of class
participation; (2) social interaction, which analyzed pupil-teacher discourse to determine classroom climate; (3) teacher characteristics, which reviewed the qualities that are possessed by good teachers; and (4) cognitive interaction, which analyzed the levels of thinking that occur in the classroom. This present study may be classified under cognitive interaction.

Studies aimed at determining the level of cognitive behavior have found that a major portion of classroom time is spent in memory-type learning. Memory-type learning is classified as "knowledge" in Bloom's Taxonomy of Educational Objectives and is considered the lowest form of cognitive development. Gallagher (1965) stated that the majority of classroom questions is at the cognitive-memory level, while Davis and Tinsley (1967) found that eighty-seven percent of all classroom questions could be included under memory or comprehension.

**Objectives**

In this study, an effort was made to increase teacher effectiveness in questioning techniques by applying the Bloom-Sanders Taxonomy (see Appendix A) to the classroom learning situation. By making teachers aware of the different cognitive levels and encouraging them to ask questions from all levels, it was hoped that teacher effectiveness would be increased. With increase in
teacher effectiveness in questioning skill, it was felt that high-level goals may be more easily attained.

There were three main objectives associated with this study. The first objective was to determine if student-teachers could achieve higher cognitive classroom behavior when they applied the *Taxonomy* to their questioning techniques in the classroom. The second objective was to determine if three investigators, who had acquired considerable knowledge of the *Taxonomy* could, through the use of a Teacher Pupil Question Inventory (TPQI) based on the *Taxonomy*, agree upon the level of cognitive behavior occurring in the teacher-pupil dialogue within the classroom. The third objective was to determine if regular classroom teachers could, through the use of the TPQI, agree upon the level of cognitive behavior occurring within the classroom during the teacher-pupil dialogue.

Answers were also sought to the following questions:

(1) Is there a difference in the level of cognitive classroom behavior achieved among student-teachers who have been instructed in the use of the *Taxonomy*?

(2) Is there an increase over time in the level of cognitive classroom behavior of student-teachers who have been instructed in the use of the
(3) Is there a difference in the mean rating scores of investigators \((R_1-R_3)\) and the mean ratings scores of co-operating teachers \((R_1-R_3)\)?

(4) Is there an increase over time in the agreement of co-operating teachers' ratings of student-teachers?

(5) Is there a difference in the percentage of memory questions asked by a group of student-teachers who had training in the use of the Taxonomy compared with (1) a control group of student-teachers who have not had instruction in the use of the Taxonomy, and (2) a group of student-teachers, previously cited in the literature (Davis and Tinsley, 1967), who have not had instruction in the use of the Taxonomy?

**Procedures**

The basic approach was to instruct a group of student-teachers in the classroom use of the Taxonomy (experimental group) and compare it with a group of student-teachers not instructed in the use of the Taxonomy (control group) but given equal instruction in other elements of pedagogy. Tape recordings were made of lessons taught by student-teachers of both groups.
The investigators \((R_1-R_3)\) rated each of these taped lessons. The TPQI, which employed a point score scale, was used to determine the level of achievement. The statistical method of analysis of variance was used to determine whether there was a significant difference (1) between the two treatments, and (2) among the three investigators.

At the same time, those co-operating teachers working with the experimental group of student-teachers were given instruction in the use of the Taxonomy. The co-operating teachers \((R_4-R_9)\) then rated the tapes made by each student-teacher of the experimental group, using the TPQI. The ratings of the co-operating teachers were compared with one another by analysis of variance to determine if regular classroom teachers can agree in their rating of student-teachers.

**Design**

To obtain answers to the first two objectives, a four factor design with repeated measures was used. The factors included (1) the treatment variable of whether the student-teachers had training or no training in the use of the Taxonomy, (2) the student-teacher variable, (3) the rater (investigators) variable, and (4) the trial variable.
To obtain an answer to the third objective, a three factor design with repeated measures was used. The factors included (1) the student-teacher variable, (2) the rater (three investigators and six co-operating teachers) variable, and (3) the trial variable.

The complexity of the two designs made it possible to obtain answers to four of the other questions put forward in this study. Answers to the fifth question were obtained by performing chi-square tests on the percentages of memory and above-memory questions asked by the different groups of student teachers. The experimental group percentages were compared with those of (1) the control group, and (2) the Davis-Tinsley group.

Findings

Test results relating to the three main objectives were:

(1) There was a significant difference (p<.01) between treatments. (The experimental group achieved a higher level of cognitive classroom behavior).

(2) There was no significant difference among the rating scores of investigators.

(3) There was a significant difference (p<.01)
among the rating scores of the co-operating teachers. (It was found, however, that five of the six co-operating teachers agreed in their rating of student teachers).

Other findings:

(4) There was a significant difference \((p<.01)\) in the level of cognitive classroom behavior achieved among student-teachers who had instruction in the use of the Taxonomy.

(5) There was no increase over time in the level of cognitive classroom behavior of student-teachers who had been instructed in the use of the Taxonomy.

(6) There was a significant difference \((p<.01)\) between the mean scores of the investigators and the mean scores of the co-operating teachers. (It was found, however, that five of the six co-operating teachers' scores did not differ from the investigators' scores. It can be stated, therefore, that eight of the nine raters agreed in their rating of student-teachers).

(7) There was no increase in rating agreement of co-operating teachers over time.

(8) There was a significant difference \((p<.05)\) between the percentage of memory questions asked
by the experimental group and the control group of student-teachers. (Results favored the experimental group).

(9) There was a significant difference ($p < .001$) between the percentages of memory questions asked by the experimental group and the Davis-Tinsley group. (Results favored the experimental group).

**Conclusions**

A number of conclusions can be drawn from this study. First, instruction in the use of the *Taxonomy* as a tool for clearly identifying the different levels of intellectual behavior will be helpful to both student-teachers and experienced teachers alike. Second, since the TPQI is simple to use and appears to have high reliability, it may be a useful device for supervisors of teachers. Third, the close agreement by raters on the cognitive behavior occurring in the classroom suggests that the *Taxonomy* can serve as a common educational language for the improvement of instruction.

Finally, it must be conceded that social study goals which call for higher levels of thinking are not being achieved. The findings reported here are similar to those of Gallagher, and Davis and Tinsley. Together, they challenge the time-honored practice of asking questions as a
method of teaching. A corollary to the above is that teachers are not being given the proper training to successfully implement the higher levels of thinking in their classrooms.

Recommendations

The findings reported herein appear to suggest a promising approach toward the improvement of teacher education and pupil achievement. It appears desirable to include study of the Taxonomy in the teacher-training curriculum as a tool for raising the level of classroom learning. The following recommendations are suggested for introducing the Taxonomy to teachers in a meaningful way: (1) combine the teaching of the Taxonomy with Taba's (1967) schema of cognitive tasks; (2) use the Taxonomy with micro-teaching.

The success of the TPQI as a measurement of cognitive behavior suggests that the Taxonomy could be used with other measuring instruments. The Taxonomy might be used to add a cognitive dimension to the OSCAR, or it might be employed with Flanders' interaction analysis to determine the effect of teaching style on cognitive achievement.

In view of the small sample size, this study should be replicated on a larger scale to verify the findings
and conclusions presented here. A suggested modification might be to teach the Taxonomy in three phases. Phase one would concentrate on encouraging teachers to ask above-memory questions. Phase two would emphasize convergent thinking (application questions). Phase three would emphasize divergent thinking (analysis, synthesis, and evaluation questions).

Several other questions might also be tested in subsequent studies. For example: (1) Will the cognitive response level increase if the pupil is taught to become aware of the level of his expected response? (2) Does eliciting longer answers from the pupil tend to raise the cognitive level of the response? (3) Is there a correlation between the type of teacher (according to the Flanders model) and the cognitive level attained in the classroom (as measured by the Taxonomy)? (4) Do students operate at the same cognitive level in one subject as they do in another?
APPENDIX
APPENDIX A

THE BLOOM-SANDERS' TAXONOMY OF EDUCATIONAL OBJECTIVES: COGNITIVE DOMAIN

Memory - ability to recognize or recall information presented earlier. It includes definitions, generalizations, and values.

Translation - ability to express information in a different form. A student translates when he expresses something in his own words. He also translates when he explains what he sees in a picture or draws a picture of something that he heard or read about. Acting out of historical events is another form of translation.

Interpretation - ability to perceive a relationship between two ideas. This relationship is perceived by the student through use of his own common sense.

Questions which may be placed at the interpretation level tell the student explicitly what to do and have one or only a few logical answers. (e.g. Compare the Northern and Southern view on slavery prior to 1960.) The different kinds of relationships include comparison, implication (informal deduction), generalization (informal induction), value, skill of definition, quantity, and cause and effect.

Note: If, in making comparisons, the student employs a more formal approach, using the rules of logic, he is operating at the analysis level.

Application - ability to use ideas, principles, and generalizations in new situations.

Questions are considered to be at the application level if they have problem-solving power, deal with the whole of ideas and skills, and include a minimum of directions (as the student is expected to know what to do). (e.g. In 1860, how did the North and South differ?) This question is much broader than the example used in the interpretation category.
APPENDIX A – Continued

Analysis - ability to break down material into its parts by employing the formal rules of logic. Since very little teaching time is given to formal instruction in the parts and processes of reasoning (induction, fallacies, deduction, and semantics), very few analysis questions can be employed in the average classroom.

Note: Sanders does suggest that analysis questions related to developing generalizations by the inductive approach may be used with elementary school children.

Synthesis - ability to draw elements from many sources to form a unified structure not clearly perceived before. Two important characteristics of synthesis questions are that they allow the student great freedom in seeking a solution, and that their solution requires a product in the form of a unique plan, a communication, or a set of abstract relations.

Evaluation - ability to make a judgment about the value of ideas, solutions, methods, materials, etc., using criteria developed by the student, himself (not by the teacher).

Evaluation questions require the student to perform two operations. First, he must establish appropriate standards or values. Second, he must determine how closely the idea or object meets these standards.

It is important to remember that evaluation questions deal with values and not with facts or opinions. They are always somewhat subjective because either their standard cannot be proved to be correct, or the idea to be judged cannot be proved to violate the standard.

Note: Questions in which the teacher specifies the values for making a judgment fall under the interpretation category.
APPENDIX B

THE FLANDERS-AMIDON MODEL OF CATEGORIES FOR INTERACTION ANALYSIS

1. **Accepts Feeling**: Accepts and clarifies the feeling or tone of the students in a non-threatening manner. Feelings may be positive or negative. Predicting and recalling feelings are included.

2. **Praises or Encourages**: Praises or encourages action or behavior. Jokes that release tension, not at the expense of another individual, nodding head or saying "uh huh" or "go on" are included.

3. **Accepts or Uses Ideas of Students**: Clarifies, builds, or develops ideas or suggestions made by the student. (As teacher brings more of his own ideas into play, shift to category five.)

4. **Asks Questions**: Asking a question about content or procedure with the intent that a student answer.

5. **Lectures**: Giving facts or opinions about content or procedure; expressing own ideas, asking rhetorical questions.

6. **Gives Directions**: Expresses procedures, commands, or orders with which the student is expected to comply.

7. **Criticizes or Justifies**: Statements that are intended to change student behavior from non-acceptable to acceptable pattern; bawling someone out, stating why the teacher is doing what he is doing, extreme self-reference.

8. **Student Talk - Responses**: Talk by students
in response to teacher. Teacher initiates the contact or solicits student statement.

9. **Student Talk - Initiation:** Talk by students which they initiate. (If "calling on" student is only to indicate who may talk next, observer must decide whether student wanted to talk. If he did, use this category.)

10. **Silence or Confusion:** Pauses, short periods of silence, and periods of confusion in which communication cannot be understood by the observer.

### The Flanders' Matrix

<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>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></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>
</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>
</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>
</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>
</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>
</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>
</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>
</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>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>%-age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX C

WEEKLY PROCEDURES

Co-operating Teachers

First Week

Group Meeting

Introduce Taxonomy. Hand out paper on definitions (APPEN. A)
Use some selected classroom questions to illustrate how to determine their cognitive level.

Second Week

Group Meeting

Introduce the TPQI. Give teachers some practice in using the TPQI by playing a portion of a previous taped lesson for them to rate.

Experimental Group of Student-Teachers

Group Meeting

Explain Interaction Analysis. Introduce the Taxonomy (Same as Col. 1). Ask them to tape record one social studies lesson each week so that it might be analyzed, using the Taxonomy. Provide them with tapes.

Control Group of Student-Teachers

Group Meeting

Explain Interaction Analysis. Introduce the Flanders-Amidon Model (APPEN. B) Ask them to tape record one social studies lesson each week so that it might be analyzed, using the Flanders-Amidon Model. Provide them with tapes.

Group Meeting

Meet individually with each and listen informally to a portion of their taped lesson. Emphasis will be on technical quality. The Taxonomy will only be briefly discussed.

Meet individually with each and listen informally to a portion of the taped lesson. Emphasis will be on technical quality. The Flanders-Amidon Model will only be briefly discussed.
### Co-operating Teachers

**Third Week**

**Group Meeting**

Listen to and rate tapes of six student-teachers of experimental group.

### Experimental Group of Student-Teachers

- Meet individually with each. A second taped lesson will be provided by her. The student-teacher will be asked to state her objectives and the procedures which she followed in conducting this lesson. A portion of the tape will be analyzed by the student-teacher and the investigator to determine the level of cognitive behavior that had been achieved. *(Provide each with tape #2.)*

### Control Group of Student-Teachers

- Meet individually with each. A second taped lesson will be provided by her. The student-teacher will be asked to state her objectives and the procedures which she followed in conducting this lesson. A portion of the tape will be analyzed by the student-teacher and the investigator to determine the type of social interaction that was going on. *(Provide each with tape #2.)*

### Fourth Week

- Meet individually with each to evaluate a portion of the third taped lesson of the student-teacher assigned to that co-operating teacher. Also discuss any problems that they might be having concerning the Taxonomy.

- Meet individually with student-teachers and demonstrate how one uses the TPQI in evaluating the lesson. For the following week, ask the student-teachers to evaluate their next taped social studies lesson, using the TPQI.

- Meet individually with student-teachers and demonstrate how one uses the Flanders-Amidon Model and Matrix to evaluate a lesson. For the following week ask the student-teachers to evaluate their next social studies lesson, using the Flanders Matrix.
Co-operating Teachers | Experimental Group of Student-Teachers | Control Group of Student-Teachers
---|---|---

**Fifth Week**

**Group Meeting**

Evaluate tapes of the six student-teachers of the experimental group.

Meet only long enough with each to pick up her evaluation of her own lesson as was requested on the previous week. (Provide each with tape #3.)

Meet only long enough with each to pick up her evaluation of her own lesson as was requested on the previous week. (Provide each with tape #3.)

**Sixth Week**

Meet individually with each to discuss any problems she might have with Taxonomy or the evaluation of student teachers.

Discuss the next social studies lesson to be taught by each. Include objectives and methods plus how the student-teachers might structure the lesson to achieve higher levels of cognitive behavior.

Discuss the next social studies lesson to be taught by each. Include objectives and methods that might be used. Also discuss whether the lesson would be more effective if the teacher assumed the role of a dominant type or integrative type teacher.

**Seventh Week**

**Group Meeting**

Evaluate tapes of the six student-teachers.

Will not meet with student-teachers.

Will not meet with student-teachers.
### APPENDIX D

**TEACHER-PUPIL QUESTION INVENTORY**

<table>
<thead>
<tr>
<th></th>
<th>Memory</th>
<th>Analysis</th>
<th>Rater</th>
<th>Observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Translation</td>
<td>Synthesis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Interpretation</td>
<td>Evaluation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Application</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject</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>Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher</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>Pupil</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>Teacher</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>Pupil</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>Teacher</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>Pupil</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>Teacher</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>Pupil</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>Teacher</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>Pupil</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>Teacher</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>Pupil</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>
Individual composite scores for both the experimental and control groups of student-teachers have been compiled into a matrix and are presented in Appendix E. Each number represents the raw score which a student-teacher received from an individual rater for one of the three trial periods. The student-teacher's total score for each trial period is given at the right (outside the box) of the individual scores. Similarly, individual rater scores appear in the columns and their totals are shown at the bottom (below the box).

It should be remembered that only the investigators ($R_1-R_3$) evaluated the student-teachers in the control group. Consequently, there are only one-third as many scores for the control group as compared with the experimental group. This, in no way, affects the results as the experiment was designed with this in mind. In order that the reader may make a quick comparison of groups, the investigators' ($R_1-R_3$) scores for the experimental group are given at the far left while the control group scores are shown at the far right.

In Appendix F, the actual TPQI scores for all student-teachers are reproduced. The individual rater scores,
involving the ten questions and ten responses associated with each taped lesson, are given here. The total score, for each TPQI, is shown at the right. This total score was used in compiling the data in Appendix E.
# Appendix E

## Composite Scores

### Experimental Group

<table>
<thead>
<tr>
<th>Totals</th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
<th>R5</th>
<th>R6</th>
<th>R7</th>
<th>R8</th>
<th>R9</th>
</tr>
</thead>
<tbody>
<tr>
<td>136</td>
<td>S1</td>
<td>49</td>
<td>49</td>
<td>38</td>
<td>53</td>
<td>39</td>
<td>52</td>
<td>36</td>
<td>52</td>
</tr>
<tr>
<td>112</td>
<td>S2</td>
<td>38</td>
<td>40</td>
<td>34</td>
<td>37</td>
<td>38</td>
<td>42</td>
<td>38</td>
<td>45</td>
</tr>
<tr>
<td>136</td>
<td>S3</td>
<td>44</td>
<td>50</td>
<td>42</td>
<td>45</td>
<td>43</td>
<td>53</td>
<td>63</td>
<td>45</td>
</tr>
<tr>
<td>74</td>
<td>S4</td>
<td>26</td>
<td>24</td>
<td>24</td>
<td>26</td>
<td>26</td>
<td>29</td>
<td>29</td>
<td>32</td>
</tr>
<tr>
<td>144</td>
<td>S5</td>
<td>45</td>
<td>50</td>
<td>49</td>
<td>43</td>
<td>49</td>
<td>46</td>
<td>39</td>
<td>43</td>
</tr>
<tr>
<td>124</td>
<td>S6</td>
<td>38</td>
<td>41</td>
<td>45</td>
<td>32</td>
<td>33</td>
<td>44</td>
<td>47</td>
<td>29</td>
</tr>
<tr>
<td>726</td>
<td></td>
<td>240</td>
<td>254</td>
<td>232</td>
<td>236</td>
<td>228</td>
<td>260</td>
<td>272</td>
<td>234</td>
</tr>
</tbody>
</table>

### Control Group

<table>
<thead>
<tr>
<th>Totals</th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
<th>R5</th>
<th>R6</th>
<th>R7</th>
<th>R8</th>
<th>R9</th>
</tr>
</thead>
<tbody>
<tr>
<td>128</td>
<td>S1</td>
<td>42</td>
<td>42</td>
<td>44</td>
<td>40</td>
<td>40</td>
<td>53</td>
<td>55</td>
<td>43</td>
</tr>
<tr>
<td>108</td>
<td>S2</td>
<td>40</td>
<td>32</td>
<td>36</td>
<td>35</td>
<td>30</td>
<td>30</td>
<td>37</td>
<td>35</td>
</tr>
<tr>
<td>129</td>
<td>S3</td>
<td>43</td>
<td>39</td>
<td>47</td>
<td>48</td>
<td>54</td>
<td>60</td>
<td>50</td>
<td>47</td>
</tr>
<tr>
<td>119</td>
<td>S4</td>
<td>42</td>
<td>40</td>
<td>37</td>
<td>39</td>
<td>40</td>
<td>41</td>
<td>46</td>
<td>48</td>
</tr>
<tr>
<td>141</td>
<td>S5</td>
<td>48</td>
<td>42</td>
<td>51</td>
<td>47</td>
<td>41</td>
<td>46</td>
<td>64</td>
<td>48</td>
</tr>
<tr>
<td>134</td>
<td>S6</td>
<td>48</td>
<td>40</td>
<td>46</td>
<td>52</td>
<td>60</td>
<td>50</td>
<td>62</td>
<td>61</td>
</tr>
<tr>
<td>759</td>
<td></td>
<td>263</td>
<td>235</td>
<td>261</td>
<td>261</td>
<td>265</td>
<td>280</td>
<td>314</td>
<td>282</td>
</tr>
</tbody>
</table>

### White Group

<table>
<thead>
<tr>
<th>Totals</th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
<th>R5</th>
<th>R6</th>
<th>R7</th>
<th>R8</th>
<th>R9</th>
</tr>
</thead>
<tbody>
<tr>
<td>127</td>
<td>S1</td>
<td>42</td>
<td>43</td>
<td>42</td>
<td>41</td>
<td>37</td>
<td>38</td>
<td>35</td>
<td>38</td>
</tr>
<tr>
<td>150</td>
<td>S2</td>
<td>50</td>
<td>52</td>
<td>48</td>
<td>36</td>
<td>33</td>
<td>36</td>
<td>40</td>
<td>48</td>
</tr>
<tr>
<td>149</td>
<td>S3</td>
<td>52</td>
<td>50</td>
<td>47</td>
<td>50</td>
<td>52</td>
<td>53</td>
<td>56</td>
<td>55</td>
</tr>
<tr>
<td>116</td>
<td>S4</td>
<td>44</td>
<td>35</td>
<td>37</td>
<td>41</td>
<td>42</td>
<td>40</td>
<td>43</td>
<td>46</td>
</tr>
<tr>
<td>147</td>
<td>S5</td>
<td>54</td>
<td>45</td>
<td>48</td>
<td>41</td>
<td>41</td>
<td>57</td>
<td>51</td>
<td>60</td>
</tr>
<tr>
<td>135</td>
<td>S6</td>
<td>45</td>
<td>46</td>
<td>44</td>
<td>34</td>
<td>40</td>
<td>42</td>
<td>44</td>
<td>41</td>
</tr>
<tr>
<td>824</td>
<td></td>
<td>287</td>
<td>271</td>
<td>266</td>
<td>243</td>
<td>245</td>
<td>250</td>
<td>275</td>
<td>279</td>
</tr>
<tr>
<td>S4</td>
<td>S1</td>
<td>S2</td>
<td>S3</td>
<td>S4</td>
<td>S5</td>
<td>S6</td>
<td>S7</td>
<td>S8</td>
<td>S9</td>
</tr>
<tr>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>2131334463</td>
<td>112111111221</td>
<td>1112212222</td>
<td>1111122334</td>
<td>1132222222</td>
<td>1122222222</td>
<td>1111111111</td>
<td>1111111111</td>
<td>1111111111</td>
<td>1111111111</td>
</tr>
<tr>
<td>3223221432</td>
<td>3212111111</td>
<td>3212111111</td>
<td>3212111111</td>
<td>3212111111</td>
<td>3212111111</td>
<td>3212111111</td>
<td>3212111111</td>
<td>3212111111</td>
<td></td>
</tr>
<tr>
<td>2333221322</td>
<td>2333221321</td>
<td>2333221321</td>
<td>2333221321</td>
<td>2333221321</td>
<td>2333221321</td>
<td>2333221321</td>
<td>2333221321</td>
<td>2333221321</td>
<td></td>
</tr>
<tr>
<td>2233221321</td>
<td>2233221321</td>
<td>2233221321</td>
<td>2233221321</td>
<td>2233221321</td>
<td>2233221321</td>
<td>2233221321</td>
<td>2233221321</td>
<td>2233221321</td>
<td></td>
</tr>
</tbody>
</table>

**APPENDIX Y-Continued**

**Experimental Group**

<table>
<thead>
<tr>
<th>S4</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
<th>S6</th>
<th>S7</th>
<th>S8</th>
<th>S9</th>
</tr>
</thead>
<tbody>
<tr>
<td>2131334463</td>
<td>112111111221</td>
<td>1112212222</td>
<td>1111122334</td>
<td>1132222222</td>
<td>1122222222</td>
<td>1111111111</td>
<td>1111111111</td>
<td>1111111111</td>
<td>1111111111</td>
</tr>
<tr>
<td>3223221432</td>
<td>3212111111</td>
<td>3212111111</td>
<td>3212111111</td>
<td>3212111111</td>
<td>3212111111</td>
<td>3212111111</td>
<td>3212111111</td>
<td>3212111111</td>
<td></td>
</tr>
<tr>
<td>2333221322</td>
<td>2333221321</td>
<td>2333221321</td>
<td>2333221321</td>
<td>2333221321</td>
<td>2333221321</td>
<td>2333221321</td>
<td>2333221321</td>
<td>2333221321</td>
<td></td>
</tr>
<tr>
<td>2233221321</td>
<td>2233221321</td>
<td>2233221321</td>
<td>2233221321</td>
<td>2233221321</td>
<td>2233221321</td>
<td>2233221321</td>
<td>2233221321</td>
<td>2233221321</td>
<td></td>
</tr>
<tr>
<td>Column 1</td>
<td>Column 2</td>
<td>Column 3</td>
<td>Column 4</td>
<td>Column 5</td>
<td>Column 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: The table continues in a similar format.*


Aschner, Mary; J.M. "Asking Questions to Trigger Thinking." National Education Association Journal, 56 (September, 1961, 44-45.


Fisher, Mary S. "Language Patterns of Pre-School Children." *Journal of Experimental Education,* 1 (September, 1932), 70-85.


Gallagher, James J. "Expressive Thought by Gifted Children in the Classroom." Elementary English, 42 (May, 1965), 559-568.


Hoenstine, Earl S. "Changes During the Student Teaching Experience in the Thinking of Student Teachers about Teaching Practices, Personality, Development, and Learning." Dissertation Abstracts, 17 (December, 1959), 2929-2930.


Klebaner, Ruth P. "Questions That Teach." Grade Teacher, 61 (March, 1964), 76-77.


Pfeiffer, Isobel, and Davis, O. L. "Teacher-Made Examinations: What Kind of Thinking Do They Demand?" NASSP Bulletin, 49 (September, 1965), 1-10.


Stevens, Romiett. The Question as a Measure of Efficiency in Instruction. New York: Teachers College, Columbia University, 1912, pp. 95.


Taba, Hilda, and Elzey, Freeman F. "Teaching Strategies and the Thought Processes." Teachers College Record, 45 (March, 1964), 524-534.


Taba, Hilda; Levine, Samuel; and Elzey, Freeman F. Thinking in Elementary School Children. San Francisco: San Francisco State College (U. S. Office of Education Cooperative Research Project No. 1574), 1964, pp. 207.

Thomas, Dorothy S.; Loomis, Alice M.; and Arrington, Ruth E. Observational Studies of Social Behavior. New Haven: Yale University, 1933, pp. 271.


