THE DEVELOPMENT AND EVALUATION OF AN OBSERVATIONAL SYSTEM DESIGNED TO MEASURE SCIENCE CLASSROOM ACTIVITY FROM THE HALL

By

NEAL JAY HOLMES

Bachelor of Science
Northeast Missouri State College
Kirksville, Missouri
1957

Master of Arts
Washington University
St. Louis, Missouri
1962

Submitted to the faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the degree of
DOCTOR OF EDUCATION
May, 1967

THE DEVELOPMENT AND EVALUATION OF AN
OBSERVATIONAL SYSTEM DESIGNED TO
MEASURE SCIENCE CLASSROOM
ACTIVITY FROM THE HALL

Thesis Approved:

Kenneth E. Wiggins
Thesis Adviser

Robert C. Fite

L. Herbert Bruneau

V. L. Pollak

N. N. Durham
Dean of the Graduate College

PREFACE AND ACKNOWLEDGMENTS

The purpose of the study was to investigate the feasibility of securing observational information relating to classroom activities in junior high school science classes by observing from the hallway.

A number of persons are due an expression of appreciation; Dr. Kenneth E. Wiggins, chairman of my committee, Drs. Jacob W. Blankenship, L. Herbert Bruneau, Robert C. Fite, and V. L. Pollock, committee members, Mr. C. W. Farnham, Superintendent of Parkway Schools, and the administrative staff, teachers and parents who participated with interest and enthusiasm.

Special thanks are due the St. Louis County Audio-Visual Library for their helpful assistance with equipment and the technical assistance and advice of Mr. John Madouras.

The writer is especially indebted to the National Science Foundation and to the administration of the Oklahoma State University for making possible the opportunity for advanced study. Special thanks go to the writer's family and close friends who have continually given support and encouragement during the course of the study.

TABLE OF CONTENTS

LIST OF TABLES

Table		Page

CHAPTER I

INTRODUCTION

The classroom observation has always been an important means of learning about the process of education. Some of the studies have involved detailed observations of one child, while other studies have searched for very specific individual and group behaviors (1). Flanders, Wright, Sears, and others (2, 3, 4, 5), have observed interaction between teachers and students by using category and sign systems of verbal interaction. Other recent studies have been conducted using 8 millimeter film, television, and various electronic devices (1, 6, 7). The system used in this study is a modified form of the Variety Section of the Cornell System (8) with the observer viewing the classroom activity from the hallway and recording by a taxonomical key.

In todays modern schools many people pass up and down the hallways while classes are in session. An observer walking down the hallway can observe behavior taking place in the observable part of the classroom. Although limited by the scope and duration of such an observation, it may be possible to answer many perplexing questions by such a technique.

The present study was made to determine the reliability, validity, and observer agreement for filmed classroom scenes photographed from the hallway. The photographed behavior consisted of thirty second scenes from junior and senior high school science classes with superimposed beep tones defining specific scenes to classify. Observers with various background training and specialized observation training were involved in this study to determine the optimum general and special training for observers.

The Need for the Study

The behavior of teachers and students in science classes is not well known. Fletcher C. Watson (9) writing on research in science teaching states:

> Research on the relations between the behavior of science teachers and other variables, such as behaviors of their pupils, is meager. The scarcity of such research on science teaching is especially unfortunate, for the structure of science and its continuous contact with mainpulatable objects offer numerous opportunities for clear and diversified appraisal of pupil behaviors. This lack of research also seems inconsistent with the numerous 'grand' objectives of science education spelled out from time to time. Achievement of such objectives does appear to be amenable to operational definition, and therefore to experimental study as a function of teacher behavior. While selection of instructional objectives may be distinguished from the instructional act itself, it does result in stressing certain pupil behavior, so that even this selection may be regarded as an act of the teacher.

Several studies of recent date give indications of mounting interest in observing science classrooms (10, 11, 12). The number of studies and the size of the observed sample is in all cases small, for example, La Shier (10) studied ten classes, Molchen (11) is working with a sample of 27, and Mathews (12) is using 18 student teachers and their supervisors. The limited numbers seem to be a result of the approach to classroom observation. Medley and Mitzel (1) summarized the problems of classroom observation as:

> The fact that observations of classroom behavior are so seldom included in investigations in which they should play a central role is easily explained. Research workers point out that observations are expensive in terms of time, money, and the professional skill demanded of observers; that observations

> constitute an invasion of privacy that teachers and administrators resent and resist (although their pupils do not); that the presence of an observer in a classroom is so disturbing that the behavior seen cannot be regarded as typical of the behavior which goes on when an observer is not present; and, above all, that most studies in the past which have employed classroom visitation have not been successful in increasing our knowledge about teaching and learning anyhow, finally they point out that the number of classrooms that is economically feasible to include in an investigation becomes so small if it is necessary to visit each one that it is better to omit observation and study a larger sample.

If an observational system can be developed that will provide reliable observational data, even in a limited way, and avoid, or reduce, the great majority of the objections mentioned by Medley and Mitzel, it would provide a worthwhile contribution to an area of expressed need. The observational system proposed in this study would eliminate many of these objections to observations by placing the observer outside the classroom and concentrating on many short time duration observations.

Leake (13) concluded in his study of teacher attitudes: "Some valid criteria for desirable learning behavior must be identified, then observational techniques developed so that the observer does not interact in any way with the observed."

The system under consideration in this study is an effort to provide a beginning in the development of such a technique.

The Scope of the Study

Observation of classroom behavior has generally been conducted by a trained observer using a category or sign system (1). These systems are designed to sample various aspects of behavior and environment. The observer records by marks or tallies, often in some rhythmic sequence. These recorded observations are subjected to analysis at a later time. The system under study is a taxonomy developed by modifying the Variety Section of the Cornell Code Digest System (8). Taxonomies have proven beneficial in situations requiring precision of separation and identification, most recently in the behavioral and learning objectives phase of education (14, 15).

The problem involves the evaluation of a taxonomical system of sampling the classroom activity by viewing the classroom from the hall. Three different occupational groups were studied to determine the optimum observer type and his or her training.

The primary purpose of the study was to develop and test an instrument modified for use in the observation of science classrooms from the hall. The study addresses itself to the following questions:

1. How reliable is an observation made from the hall?
2. How valid is such an observation using the proposed taxonomy?
3. Can observers agree on the classification of a scene viewed from the hall?
4. Is formal training in the use of the instrument helpful in increasing the reliability, validity, and coefficients of observer agreement?
5. Is formal college training in education necessary for observers?
6. Is formal college training in science and education essential for maximum reliability, validity, and observer agreement?

The procedures involved in an attempt to answer these questions were: (1) the modification of the Variety Section of the Code Digest, developed by Cornell, Lindvall, and Saupe, into a taxonomy of activity;

(2) the writing of and photographs for a manual of instruction for observers; and (3) the production of two 16 millimeter sound films of science classes as viewed from the hall. Each film consists of thirty different scenes from ten classrooms and require about fifteen minutes viewing time. These films were used as behaviors for evaluation purposes. (4) The training of thirty-six observers, eighteen will have formal training and eighteen will have only the written material and pictures in the observers manual. The observers will equally represent secondary science teachers, elementary teachers, and housewives. The details concerning the taxonomy, manual, photographs, films, observer training, and the statistical treatment will be found in Chapters III, IV, and the Appendices.

Hypotheses

The research hypotheses of this study in the null form are:

1. The previous training and background of observers do not alter the reliability of the observations using the taxonomy of classroom activity.
2. The previous training and background of observers do not alter the coefficient of observer agreement.
3. The training of observers in the use of the taxonomy of classroom activity by formal instruction does not alter the reliability of the observations.
4. The training of observers in the use of the taxonomy of classroom activity by formal instruction does not alter the coefficient of observer agreement.

Analysis of variance was used to determine the reliability and coefficient of observer agreement for each occupational group and training level. A second analysis of variance was used for analysis of observer qualities. The level of significance for rejection of the null hypothesis was set at the .05 level.

Assumptions of the Study

1. Observers not undergoing formal training do read and study the manual for the taxonomy.
2. The behavior filmed represents the true spectrum of classroom variety and was not altered by the filming process.
3. That activity can be treated apart from the total complex behavior of the classroom.

Limitations

The study is limited by the population, and the variables employed. All conclusions and implications drawn from this study must be limited to these factors.

CHAPTER II

OBSERVATIONAL TECHNIQUES, THEIR VALIDITY AND RELIABILITY

This chapter is divided into two sections. The first part discusses selected studies depicting the status of observational systems as a research method. The last part discusses various efforts to establish the validity and reliability of observational techniques.

Observational Techniques

Observational techniques have been generally divided into two distinctly different types of items. Medley and Mitzel (1) described these two approaches as:

> The first approach is to limit the observation to one segment or aspect of classroom behavior, determine a convenient unit of behavior, and construct a finite set of categories into one and only one of which every unit observed can be classified. The record obtained purports to show, for each period of observation, the total number of units of behavior which occurred and the number classifiable in each category. An approach of this type will be referred to as a <u>category system</u>.
>
> The second approach is to list beforehand a number of specific acts or incidents of behavior which may or may not occur during a period of observation. The record will show which of these incidents occurred during a period of observation and, in some cases, how frequently each occurred. An approach of this type will be referred to as a <u>sign system</u>.

Recently, researchers have begun to use the taxonomy as a means of classifying behaviors and objectives. Openshaw and Cyphert (15) developed a Taxonomy for the Classification of Teacher Classroom

Behavior with five major divisions and twenty-eight subdivisions. Openshaw and Cyphert concluded their study with the following remarks:

> There is growing conviction among several investigators that to understand teaching and learning, efforts must be focused on the further illumination of the dynamics of the classroom. The procedure and approaches used by different researchers to study this problem vary widely, but at the present state of knowledge about teaching-learning, this variety is both reasonable and desirable. Currently there is insufficient data to support strong knowledge claims about teacher-learner interaction.
>
> A system of classification and description of middle-range teacher behaviors has resulted from this research endeavor. What is needed is an extension of the system to encompass more discrete behaviors and a body of descriptive data that will provide knowledge of the relationship between a specific teacher behavior and the response possibilities and probabilities of learners.

The taxonomy idea has been used successfully in the study previously quoted as well as studies involving Bloom's Taxonomies of Classroom Behavior in both the cognative and affective domain (14).

Observational systems have been applied to many and varied research problems. The studies cited in this chapter will be organized into four sections: (1) selected studies before World War II, (2) effective teacher behavior, (3) classroom climate, and (4) quantitative descriptions of teaching.

Studies Before World War II

The history of objective observational systems is a relatively short one with a probable beginning shortly before World War I. One of the earlier systems of classroom observation was developed by Horn (16) using a sign system to record recitation and requests for recitation on a pupil seating chart. This system was expanded in 1928 by Puckett (17) to include several additional behaviors. The system consisted of a series of symbols representing various situations of pupil behavior but required the judgment of observers in selecting or determining the degree of quality of answers in some cases.

Wrightstone (18) developed a rather complex system of coding pupil behavior using a combination of letters and numerals to represent various predetermined behaviors. These symbols were to be recorded beside the names of the various students as the classroom activity was in progress. While Puckett was not interested in scoring his records, Wrightstone's technique was developed specifically for measuring teacher conduct of class discussions. It is worthy to note the situations in which the Wrightstone instrument is intended for use as stated in his directions for use:

> Directions for Use: This code is to be used when a class is planning, reporting, or discussing units of work, problems, or activities. It is to be used also when a class is reciting on a lesson. It is planned to cover the major types of pupil responses in group recitations and discussions.

This instrument was not designed for the activity oriented science classroom, particularly the laboratory situation.

Jersild and Meigs (19) summarized the research methods of direct observation between 1925 and 1939 citing fifty-nine studies involving some form of direct observation. The following statement regarding the position of the observers is of interest to the present study.

> If the observer desires only a rough measure of certain characteristics or outstanding episodes, he can station himself at a strategic point and note such happenings as flow before him, without systematic attention to each member of the group.

The effect of the observer upon the observed was also noted by Jersild as a problem to those interested in the use of classroom observation systems. Two alternatives were mentioned:

> One-way vision screens or windows may be used for some purposes, but in many circumstances it would not be feasible to conceal the observer. The eaves-dropping technic can be duplicated to some degree in 'free' situations when the observer or auditor records what transpires while ostensibly he is uninterested or preoccupied with other matters.

Apparently only two studies involving the so-called eaves-dropping technic had been reported up to the 1939 summary and no additional studies have been discovered by the author. The studies mentioned were by Carlson (20) and Landis (21) relative to the conversation habits of men and women.

Carlson investigated the conversation patterns of attending men and women during the intermissions of thirteen concerts of the Minneapolis Symphony Orchestra. The investigators carried prepared cards listing eleven topics of conversation to be scored during the overheard conversations. Carlson reported a high degree of reliability from this system.

Landis and Burtt (22) used the eaves-dropping technique as early as 1924 to study five hundred fragmentary conversations overheard in Columbus, Ohio. They were successful in determining marked differences in interests of business and industrial workers as compared to students. It was also found by this method that environment of the overheard conversation had marked influence on the topic.

The use of movie pictures is not new in the development and testing of observational systems. One of the earlier users of this technique was the Yale Institute of Human Relations. Arrington (23), in 1932, reported:

> The department of social science methodology at the Yale Institute of Human Relations has recently used motion pictures to study technical problems which cannot be solved adequately

> by data from life-situations. Three observers recorded simultaneously the behavior of a given character, noting on mimeographed recording bands the occurrence of specific items within 5-second intervals. Two aspects of behavior were observed - language and physical contacts with persons. Fourteen films were viewed each a minimum of eight times. In analysis of the data, adjustment had to be made for discrepancies in the timing of events, caused by inadequately synchronized stop watches and differences in the running-speed of the same films at different times. The self-consistency of observers from observation to observation was noted. Improvement in records caused by practice and training, and the reliability of observers in recording the occurrence and in interpreting the nature of events according to predetermined definitions reached the maximum in the third of four observations.

The Yale study considered the reliability of observers based on agreement of two observers and considered three sources of discrepancies. These three sources of discrepancies were: (1) differences in seeing the events, (2) differences in interpreting the events seen, and (3) differences in timing the occurrence of events.

Willard C. Olson and Elizabeth M. Cunningham (24) reported on seventy-six studies involving observational techniques as early as 1934. None of these studies involved the eaves-dropping technique. The "time-sampling technique" under study at the time was discussed and described by Olson and Cunningham as:

> The systematic recording of a definitely delimited unit of behavior described in terms of action over a stated time interval yielding quantitative individual scores by means of repeated time units.

The time-sampling technique reported in this research showed units of time for observations from five seconds to five minutes. The time-sampling technique includes the following information:

1. Observation by an eye witness.
2. Behavior to be observed defined in terms of overt action.
3. Behavior of an individual or group observed for a stated time unit, usually short.

4. A stated number of repetitions of the time unit employed.
5. An individual score based upon the number of time units in which the defined behavior occurs.

Thorndike, Loftus, and Goldman (25) compared observed behaviors of students in four activity schools and four control schools of the New York City Public School System. The observational system consisted of seven categories (cooperative, critical, experimental, leadership, recitational, self-initiated, and negative work-spirit behaviors), and an anecdotal record. Eighty students were observed during the winter and spring of 1940. Each child was observed on twenty-five occasions for five minutes each time. Percent agreement was computed comparing the various observer records. The range of agreements was from 46% to 53%. The researchers concluded that considerable similarity was apparent between classes. The teacher control groups showed more teacher questioning and pupil answering categories.

Effective Teacher Behavior

Many studies have been designed around the teacher effectiveness versus teacher behavior criteria in an effort to develop predictable means of identifying "good" teachers. Barr (26) did one of the earlier studies in the area of social studies involving good and poor teachers. One of the items recorded on the general observation schedule of Barrs' was the various kinds of materials and equipment in the classroom.

Jayne (27) reported in 1945 on two studies done using tape recordings of 38 classrooms and counts of behavior involving 184 different categories. One hundred of these items were omitted from the study because of the infrequence of occurrence. The remaining 84 categories were reduced to eleven showing some promise of predicting outcomes from the observed behavior. The author concluded:

1. Specific, simple items did not correlate with outcomes, even though many could be reliably recorded.

2. It was possible to combine items into "indices" which did correlate with outcomes.

3. Expert's ratings had low reliability and validity.

4. Behaviors observed should be relevant to the outcomes measured.

Morsh (28) studied the behavior of U.S. Air Force Instructors and their students in an effort to answer the following questions:

1. Determining which behaviors of instructors and their students can be reliably and systematically observed

2. Determining whether or not instructor behaviors that can be observed tend to be typical and consistent

3. Determining the relationship between elements of an instructor's or student's behavior and the amount students learn or the manner in which students are graded by their supervisors

Two items are of particular interest in the Morsh study in relation to the present one. First, the observers were without previous training and experience in rating instructors, and secondly, the observers obtained data involving the verbal and non-verbal behaviors of both the instructor and student. The conclusions seemed to indicate observed student alertness and achievement are closely related.

While Morsh attempted to develop a technique suitable for non-professional observers. Wright (3) in a study involving the development of a category system for the study of behaviors in secondary school mathematics classrooms concluded that observers must be experienced in the subject matter field of mathematics and be trained by the instrument developer.

Ryan (29) has conducted rather extensive studies in the interrelatedness of observed student behavior as compared with observed teacher behavior. Ryan summarizes:

> For elementary school classes, high positive relationships were noted between observers' assessments of 'productive pupil behavior' (e.g., assessments presumed to reflect pupil alertness, participation, confidence, responsibility and self-control, initiating behavior, etc.) and observers' assessments of previously identified patterns of teacher behavior which seemed to refer to understanding, friendly class-behavior; and stimulating, original classroom behavior.
>
> For secondary school classes, low positive relationships appeared between productive pupil behavior and the above named categories of teacher behavior, with a tendency for the stimulating, original teacher classroom behavior pattern to show a slight higher correlation with pupil behavior than the understanding, friendly, or the organized, businesslike teacher behavior patterns.

Many studies have been reported attempting to relate effective teacher behavior to measured pupil growth (30). The results of these studies reveal uniformly negative results. Anderson (31) states: "No appreciable relationships exist between rating criteria and pupil attainment criteria."

Jayne (27) concluded in a study involving supervisory ratings that: "Supervisory ratings seem to lack reliability and validity as measures of pupil gain."

Jones (32) attempted to relate supervisory ratings to measured pupil gain and concluded: "Whatever pupil gain measures in relation to teaching ability it is not that emphasized in supervisory ratings." Bibliographies by Domas and Tiedman (33) and Barr (34) report over 1000 studies attempting to relate teacher personality characteristics and teacher effectiveness.

Getzels and Jackson (35) concluded:

> Despite the critical importance of the problem and half-century of prodigious research effort, very little is known for certain about the nature and measurement of teacher personality, or about the relation between teacher personality and teaching effectiveness. The regrettable fact is that many of the studies so far have not produced significant results. Many others have produced only pedestrian findings. For example, it is said after the usual inventory tabulation that good teachers are friendly, cheerful, sympathetic, and morally virtuous rather than cruel, depressed, unsympathetic, and morally deprived. But when this has been said, not very much that is especially useful has been revealed. For what conceivable human interaction -- is not the better if people involved are friendly, cheerful, sympathetic, and virtuous rather than the opposite? What is needed is not research leading to the reiteration of the self-evident, but to the discovery of specific and distinctive features of teacher personality and of the effective teacher.

Classroom Climate

Measuring classroom climate by observation has received more attention than perhaps any other facet of classroom interaction (1). Thomas (36), as early as 1929, began studies of individuals and their interactions. Four distinct patterns of studies followed. The first technique involved a single child his movements and activities, while another involved the recording of physical contacts made by children. The third technique involved stenographic records of everything said by or to the observed child, and finally observations were made of social groups.

Anderson and Brewer (37, 38, 39) developed two category systems; one for recording the "Dominative" and "Integrative" behaviors observed in teacher student contacts, and a second used concurrently with the first to record individual student behaviors. Each individual student was observed in five-minute intervals until twenty-four observations had been made (about two hours). The number of classrooms was limited by the very nature of the studies and reliabilities in comparison to various classrooms and teachers were limited.

Withall (40, 41) developed a system of categories for use in the analysis of typewritten transcripts taken from sound tapes of classroom behaviors. He demonstrated that the verbal statements of teachers could be classified into categories for measurement and analysis. Withall developed a set of seven categories, similar in nature to Anderson's (42) Dominative-Integrative rating, called the "Social-Emotional Climate Index." It was comprised of criteria whereby teacher statements were distributed as follows: (1) learner-supportive statements or questions, (2) acceptant or clarifying statements or questions, (3) problem-structuring statements or questions, (4) neutral statements evidencing no supportive intent, (5) directive statements or questions, (6) reproving disapproving or disparaging statements or questions, and (7) teacher-supportive statements or questions.

Hughes and her associates (43) developed a comprehensive set of categories for the classification of teacher behaviors. This system also contained seven major categories with thirty-one subdivisions. The seven categories were: (1) controlling functions, (2) imposition, (3) facilitating, (4) content development, (5) personal response, (6) positive affectivity, and (7) negative affectivity. In the studies by Hughes, teachers were observed in half hour intervals with each teacher being observed on three different occasions. Hughes concluded that teachers' behavior patterns are stable through time; that, for example, the number of controlling acts exhibited by a given teacher in different situations does not vary significantly when compared with the interaction between situations and teachers.

The thirty-one teaching functions forming the seven major categories of the Hughes system were developed from analysis of about 1,000 written records of actual teaching of some sixty teachers in both elementary and secondary schools.

Amidon and Simon (5) established three main divisions to classify systems used to collect and categorize observational data involving teacher-pupil interaction. These three divisions were cognitive, affective, and multi-dimensional. The affective category has been identified with the Flanders system of interaction analysis, of 171 studies involving observation reported, 83 involved the Flanders Interaction System (2). Flanders discusses the meaning of the terms Interaction and Analysis, summarizing his remarks with the conclusion that although rather difficult to describe because of the diverse usage of terms in other disciplines the concurrent classroom interaction analysis describes the attempt to abstract spontaneous behavior into a set of categories. The ten categories of the Interaction Analysis System are:

1. Accepts Feeling
2. Praises or Encourages
3. Accepts or Uses Ideas of Student
4. Asks Questions
5. Lecturing
6. Giving Directions
7. Criticizing or Justifying Authority
8. Student Talk - Response

9. Student Talk - Initiation

10. Silence or Confusion

Observers classify verbal statements every three seconds. The observer's record consists of a series of numerals representing the categories and the order of occurrence of these behaviors. Pairs of these numerals make up the coordinates for a matrix plot of the observed spontaneous behaviors.

La Shier (10) reported in 1966 on a study involving eighth grade students using a modified laboratory block program of the Biological Sciences Curriculum Study. Student teachers were observed using the Flanders system. The study indicated significant relationships between the verbal influence of the student teachers, the freedom of participation of the pupils, and the subsequent achievement of constructive attitudes of the students. The comparison of the two matrices indicated that the verbal patterns of the indirect group of student teachers differed substantially from that of the direct group. The indirect group was more accepting of student-initiated ideas, tended to encourage these ideas more, and also made more of an effort to build upon these ideas than did the direct group of student teachers. The indirect group also spent less time lecturing, and giving directions.

Parakh (44) developed an observational system involving forty-five categories and tried it out on ten biology teachers. A conspicuous feature was the preponderance of teacher talk. The average or composite teacher talked about 75 percent of the total time in lecture-discussion classes and about 50 percent of the time in laboratory situations. Parakh pointed out the need for additional quantitative information about the manner in which science materials are implemented into the classroom.

Quantitative Description of Teaching

Other studies have made an attempt to measure teacher classroom behavior as such, to describe in quantitative terms as much as possible of what goes on in the classroom without reference to the relationship of behavior to teacher effectiveness or to any psychological theory. These studies have relied heavily upon those of classroom climate but have attempted to measure differences in classrooms without regard to the effectiveness component.

Cornell, Lindvall, and Saupe (8) developed an instrument for the expressed purpose of "measuring differences in classrooms as a means of characterizing differences in school systems." The instrument, "Classroom Code Digest," consists of eight major areas with divisions within each area ranging from five in the content section to twenty-three in the variety section. The eight major areas are:

A. Differentiation - The extent to which provision is made for individual differences among students.

B. Social Organization - The type of group structure and the pattern of interaction among individuals.

C. Initiative - The extent to which pupils are permitted to control the learning situation.

D. Content - The source and the organization of the content of learning.

E. Variety - The extent to which a variety of activities or techniques are used.

F. Competency - Differences in the technical performances of teachers.

G. Classroom Climate - Social emotional climate as it is reflected in the behavior of the teacher

H. Classroom Climate - Social emotional climate as it is reflected in the behavior of the pupils

The instrument under investigation in this study was developed from the twenty-three items of variety as sources of classroom activities. The items are:

1. Teacher lectures or reads.
2. Teacher gives demonstration.
3. Teacher shows movie or slides.
4. Pupils read text at seat.
5. Pupils read other books at seat.
6. Pupils work with workbook at seat.
7. Pupils work problems (not text or workbook) at seat.
8. Pupils study materials other than books at seat.
9. Pupils draw or paint at seat.
10. Teacher questions - pupil answers.
11. Class engages in discussion.
12. Pupils give talk or report.
13. Pupils work at blackboard.
14. Pupils read aloud from book.
15. Pupils study charts, drawings, maps.
16. Pupils work experiment.
17. Pupils construct things.
18. Pupils decorate room.
19. Pupils engage in role-playing or present play.
20. Class goes on a trip.
21. Pupils go to another room to work.
22. Pupils work in small discussion groups.
23. Pupils write test.

The Cornell system has provided a source of items for the development of at least one other instrument. The OScAR (Observation Schedule and Record) was developed by Medley and Mitzel (45) by combining and revising items from the Code Digest and the Withall System into one technique. Three basic changes in design were attempted in this combining and revising. They were:

> The first change was designed to increase observer accuracy by reducing the difficulty of the judgments required, along the line just pointed out. If an observational technique can be used successfully only by a highly trained observer, it has limited usefulness, and the results of future

> measurements made with it may be suspect because the observers may not have been adequately trained. For this reason, the categories of both Cornell and Withall were redefined in somewhat simpler terms.
>
> Experience with these two techniques also showed that the often-adopted practice of sending several observers into the classroom together (presumably so that one observer can record what another misses) is uneconomical. An average score based on observations made by two observers who see a teacher at different times obviously contains more information than one based on observations made by two observers who see the teacher at the same time, because it is based on a behavior sample that is twice as large. The OScAR was therefore designed to be used by a single observer visiting a classroom by himself.
>
> The third change was the separation of the process of scoring from the process of observing teacher behaviors. The OScAR was designed to permit recording of as many possible significant aspects of what goes on in the classroom as possible, regardless of their relationship to any dimension or scale.

The OScAR contains seven sections with several divisions within each section making possible seventy-two different selected behaviors or activities to check.

The OScAR system is still used rather extensively as shown by the number of studies reported by Amidon and Simon (5).

Travers, Wallen, and others (46) attempted to quantify the behaviors of teachers while teaching and determined, by use of a modified Withall system, that: (1) the most frequently occurring behavior was that of telling the pupils what to do; (2) the next most frequently occurring form of teacher behavior was that involved in a questioning process, and (3) the third most frequently occurring form of behavior outside of performing management functions was that of providing information.

Validity and Reliability

The question of validity of a classroom observation involves the problem of teacher-pupil behavior variability. The observations are samples of classroom behavior or events assumed to be typical of teachers and students. The question is: How typical are the observations? Is the sampling method used in selecting these classroom behaviors adequate to insure validity? Wandt and Ostreicher (47) studied the problem of teacher variability using two classes of children with instruction from three different junior high school teachers. They concluded:

1. Social and Emotional climate in the classrooms of observed teachers varied widely from occasion to occasion.

2. Social-Emotional climate in the classroom of three of the observed teachers varied systematically for two classes observed.

3. Initial observations were unreliable indices of the typical climate, even when the class was held constant.

4. In many cases it is impossible to secure an adequate picture of a teacher's behavior without making repeated observations of that teacher with different groups of students.

The Wandt and Ostreicher studies (47) indicate that a large sample of behavior is essential if a valid picture of the teacher's behavior is to be achieved.

The number and type of observational data necessary for validity seem to depend upon the nature of the data collected. Medley and Mitzel (48) report no error due to instability of teacher performance using the variety and teacher climate sections of the Cornell System and a minimum of six observations. The sample of classrooms involved in this study was confined to the elementary grades and various subject matter areas.

Perkins (49), Flanders (2), and Withall (40) found considerable difference in teaching strategy between elementary and junior high school teachers.

Various efforts have been made by researchers using classroom observations to show that these observations are valid. Wandt and Ostreicher (47) studied this problem, stating in part:

> The validity of generalizations made on the basis of these observational data depends, in the last analysis, on the extent to which certain fundamental assumptions made are: (a) the presence of the observer does not materially affect the behaviors observed, and (b) the behaviors observed in a teacher's class are representative of those which would have been observed had the observations been made at other times and with other classes.

The nature of the present study eliminates the first assumption completely, and provides a more economical means of collecting samples from the nine hundred hours of teaching during a teacher's school year.

Medley and Mitzel (48) stated:

> A measure is valid to the extent that differences in scores yielded by it reflect actual differences in behavior not differences in impressions made on different observers. For an observational scale to be valid for measuring behaviors, it must provide an accurate record of behaviors which actually occurred, scored in such a way that the scores are reliable.
>
> A measure is reliable to the extent that the average difference between two measurements independently obtained in the same classroom is smaller than the average difference between two measurements obtained in different classrooms.

Perkins (49) established validity in a study involving two minute observations using two different instruments, one for students, and one for teachers. Validity was determined in three ways:

> First, since the definitions of each category describe the kinds of things students and teachers do in the classroom, the instruments have face validity. Second, the tally sheets of the categorizations made by the observers accounted

for 120 seconds of student and teacher behavior for each time sample. This fact satisfied the criterion of inclusiveness of categories and provides further evidence of the validity of the instrument.

Finally, the learning-activity categories of Student Categories and the teacher-behavior and role categories of the Teacher Categories were validated against an external criterion. Three of the observers had each had more than 20 years' experience in public-school teaching or supervision. Without access to categorized data obtained by observations, each one wrote short descriptions on the 14 teachers and classrooms of the study, omitting all identifying data.

Twelve of the fourteen teachers in the study were properly identified by observers from the descriptions in writing and observational data on teaching behavior. The student observations could not be translated into descriptive terms for comparative purposes and no attempt was made to show the validity of the student part by this method.

Withall (41) established the objectivity of the Climate Index by comparing the scores of four trained judges obtained by scoring transcripts involving 184 teacher statements. The resulting percentage agreement ranged from 56 percent to 75 percent. One further measurement of the Index's objectivity was made by computing tetrachoric correlations between the categorization of each of the five judges on three transcripts.

Scott (50) pointed out one of the problems of using the technique of observer agreement.

> One commonly used 'reliability' index is simply P_A or the percentage of judgments on which coders agree, out of the total number of judgments. Unfortunately, this measure is biased in favor of dimensions with a small number of categories. By chance alone, one would expect better agreement on a two category than a five category scale.

Scott suggests the Index of Inter-coder Agreement correcting for the number of categories by removing the element of chance selection from the consideration. The resulting coefficient of observer agreement is calculated

$$\pi = \frac{P_o - P_e}{1 - P_e}$$

where P_o represents the percentage of judgments on which the two observers agree when coding the same data independently; and P_e is the percent agreement to be expected on the basis of chance. Scott concludes: "It can be roughly interpreted as the extent to which the coding reliability exceeds chance."

The use of the Pearson product-moment coefficient has been used as a means of estimating the internal consistency of a measuring instrument divided in halves (51). This procedure was developed independently by Spearman and Brown (51) for the solution of test construction problems. The use of analysis of variance was suggested by Jackson (52), Hoyt (53), and Alexander (54) for problems involving reliability estimations. Medley and Mitzel (48) developed a technique for determining the reliability of an observational system by use of the analysis of variance technique. They concluded that the analysis of variance has three distinct advantages over the correlational analysis:

1. It yields a single best estimate of reliability.

2. It supplies independent measures of the amount of error from different sources, and it provides simple exact tests of significance.

3. When only two sets of measurements are available, an estimate of reliability may be obtained by correlational analysis, but it is biased and has a larger sampling error than that obtained by analysis of variance. When more than two sets of measurements are available, no satisfactory estimate can be obtained by correlational analysis.

Medley and Mitzel (48) concluded that the validity of an observational system depended upon the fulfillment of three conditions:

1. A representative sample of the behaviors to be measured must be observed.

2. An accurate record of the observed behaviors must be obtained.

3. The records must be scored so as to faithfully reflect differences in behavior.

The first condition may be fulfilled if the observational system is constructed in such a way that random sampling of behaviors is possible. The second and third conditions depend upon the ability of observers to agree upon the classification of identical behaviors and the ability of the system to show differences between teacher's classroom climate.

Summary

The surveyed literature does not reveal any studies attempting to obtain information of a visual nature without sending the observers into classrooms. Sociological studies involving eaves-dropping technics and films are not new to observational studies. The use of various levels of specialization in observer training has been studied previously, and short time duration observations involving from one to several individuals have been reported in the literature. Efforts to show validity of observational systems have generally been co-relational with objectivity based on observer agreement and reliability, in recent studies, determined by the analysis of variance.

CHAPTER III

EXPERIMENTAL PROCEDURES AND PERSONNEL

Introduction

Films and data for this study were collected from the classrooms, teachers, and parents of the Parkway School District, Chesterfield, Missouri. Parkway is a rapidly growing suburban school district with 12,000 students enrolled in grades kindergarten through twelve. The district consists of nine elementary schools, two junior high schools, and one senior high school. The films for this study were made in the two junior high schools and teachers from all schools served as observers in the observation phase of the research.

There are six developmental phases in this project. They are:

1. Developing the taxonomy
2. Preparing the observer training manual.
3. Selection and filming of classroom scenes
4. Observer selection and training.
5. Collection of data
6. Analysis of the data

Developing the Taxonomy

Developing the taxonomy involved the selection of behavior judged to be identifiable from hallway observation. The behaviors used by Cornell (8) in the Variety Section of the Code Digest were used as basic functional elements in the construction of the Activity Taxonomy. The items making up the Code Digest Variety Section are reported on page 20 of this paper. Categories 1, 2, 3, 10, 11, 12, 14, and 15 were incorporated into the listening-viewing section. Categories 4, 5, 6, 7, 8, 13, and 23 were used in developing the reading and writing section. Categories 9, 16, 17, 18, 20, and 21 are included in section three of the taxonomy. Item 19 was omitted as it was judged to be irrelevant to the sciences. Since science classes were used in this particular study, the manipulative skills section was subdivided on the basis of apparatus in use and the size of the working group. Initial trial use by the author indicated a need for the section designated diversity. The taxonomy was reviewed by members of the National Science Foundation Second Academic Year Program at Oklahoma State University during the spring of 1966. As a result of this review and discussion, several suggested modifications have been incorporated in the present form of the taxonomy.

Preparing the Observer Training Manual

One-half of the observers in this study received no formal training. Informal training was provided by preparing a manual containing sections on taxonomies in general, the activity taxonomy, and illustrating photographs for classification and study. A copy of the manual and illustration photographs are included in Appendix A.

Selection and Filming of Classroom Scenes

Filming of exemplary behavior required the selection of teachers whose classes could be photographed by sixteen millimeter sound equipment located in the hallway. Ten of the twenty junior high school teachers were selected from the eighteen volunteers. All teachers were asked if they would be willing to cooperate in the filming process; only two declined. The two nonparticipating teachers were both new to the system and no effort was made to encourage their cooperation. Teachers were selected for the actual filming by using the table of random numbers (51).

Five teachers were chosen from each junior high school. The Central Junior High School teachers were assigned the code numbers one through five prefixed by the letter C, and the South Junior High School teachers were numbered one through five prefixed by the letter S. The schedules for the teachers selected are included in Table I. Classes filmed and the number of students enrolled in each of the classes are included in Table II.

Each faculty member selected for filming taught several sections of science daily. Two of these sections were selected for filming behavior. Scenes from the two classes (see Table II) were filmed on two nonconsecutive days. Teachers were informed of each day's filming, at least two days in advance. The camera and sound equipment were positioned behind a masonite sheet mounted on a movable platform. The shielding prevented the students from viewing the operator and made it possible to wait until the initial awareness of the camera had subsided before filming was started. The time required for classes to become accustomed to the camera was not recorded but generally it was less than five minutes. There were some instances where greater waiting time was necessary but, generally speaking, most classes rather quickly lost interest in the camera.

TABLE I

TEACHERS' SCHEDULE

Teacher	Period					
	One	Two	Three	Four	Five	Six
C-1	S-8	S-8	S-9	S-8		S-7
C-2	S-8		S-7	S-8	S-8	S-8
C-3	S-8	S-7		S-9	S-8	S-7
C-4		S-7	S-7	S-7	S-8	S-7
C-5	S-8	S-7	S-7		S-8	S-7
S-1	S-7	S-9	S-7	S-7		
S-2	S-7	S-7	S-7		S-7	S-7
S-3	S-9		S-9	S-9	S-9	S-9
S-4		S-9	S-9	S-9	S-9	S-9
S-5	S-7	S-8		S-8	S-8	S-8

Science Grade Seven (S-7), Eight (S-8), Nine (S-9)

TABLE II

TEACHING PERIODS AND GRADES FILMED

Teacher	Period Filmed	Grade Level	No. Students
C-1	4 and 6	8 and 7	28 and 31
C-2	5 and 6	8 and 8	28 and 30
C-3	4 and 5	9 and 8	30 and 27
C-4	5 and 6	8 and 7	31 and 28
C-5	5 and 6	8 and 7	29 and 32
S-1	2 and 4	9 and 7	26 and 31
S-2	4 and 5	7 and 7	30 and 29
S-3	5 and 6	9 and 9	28 and 27
S-4	5 and 6	9 and 9	26 and 25
S-5	5 and 6	8 and 8	29 and 31

Classrooms were furnished with individual student desks or two student laboratory tables. No attempt was made to control the classroom furniture variable in this study. Teachers often changed rooms and two of the teachers occupied three different rooms during the school day. Seven of the ten teachers remained in the same environment throughout the study.

The filming process consisted of a two-part operation involving the optical filming and the tape recording of the simultaneous sound. The film used Kodak Tri-X with magnetic stripping for sound recording. The initial sound track was recorded using the Sony-Matic TC-105 tape recorder with automatic record level and tape speed of 7½ inches per second. The taped sound was later transferred to the developed film by using the Kodak AV-105-M Magnetic Optical Projector. The film was processed by the Cine-Graphic Film Laboratories Inc. of St. Louis, Missouri. The camera used was the Kodak Model K-100 equipped with the Som Berthiot Pan-Cinor 1:2:8 F20a60 lens. The lens was used in the maximum wide angle position for all filming. Most classroom scenes were photographed with a lens setting of f-4.5, but on a few extremely dark days it was necessary to increase the opening to f-2.5. No special artificial lighting was used in any of the classrooms filmed.

Observer Selection and Training

Thirty-six observers were used in this study. Eighteen received formal training and practice in using the technique and instrument. Eighteen received no formal training but were given the manual and asked to study it and do the exercises contained in the picture section. Each of the eighteen observer groups were represented by six members of the junior and senior high school science department, six members of the elementary teaching staff, and six members of the parent teachers association. The junior high school teachers whose classes were used for behaviors were excluded from the observation groups.

Secondary science teacher observers were selected at random from the total staff less the personnel used in the photographic section of the study. The elementary level teacher observers were selected at random from the twenty-three participating in elementary science in-service education. The parent teacher association member observers were selected at random from a compiled list of high school graduate housewives willing to assist in the program.

Collection of Data

Each observer was shown the same behaviors for classification. These behaviors consist of 60 scenes of approximately five seconds each, six scenes for each of the ten participating teachers. Observers were required to observe a five second scene and select one of the elements of the taxonomy, recording their selection on the form supplied in the manual folder and included in Appendix B. The six members of each group viewed the film together but were not allowed to see any scene twice or discuss any scene.

Analysis of the Data

Scoring of the data consisted of assigning numbers to the thirty-seven divisions of the taxonomy ranging from one through thirty-seven. Each observer's record for each teacher were combined to determine a raw score. Three levels of raw scores were computed for each level of the taxonomy. These scores were analyzed by analysis of variance using the modified block procedure (48). Since all observers observed the same behavior, it was not necessary to use the visit error described by Medley and Mitzel (49). Variance was partitioned into a variance due to differences in teachers, a variance due to differences in observers, and a residual or instability variance. Reliability coefficients were computed using the formula number eight suggested by Medley and Mitzel (49).

$$R = \frac{\sigma_t^2}{\sigma_t^2 + \sigma_e^2}$$

Where σ_t^2 is the variance due to differences in teachers and σ_e^2 represents all other variances.

The coefficients of observer agreement were calculated using the Scott formula

$$\pi = \frac{P_o - P_e}{1 - P_e}$$

Where P_o is equal to the observer agreements and P_e is the probability of agreement by chance.

CHAPTER IV

RESULTS OF THE STUDY

Statistical Analysis of the Data

Eighteen analysis of variance were constructed from the cumulated raw scores using the randomized block technique (55). The raw scores and calculations are included in Appendix C. These eighteen analysis of variance were derived from the three levels of the taxonomy and the six observer teams. The three levels are designated: I for the most discriminate thirty-seven level of the taxonomy, II for the intermediate twelve classes, and III for the four class level.

The variance due to differences in teachers was compared with the sum of all other variances by use of the F test. The results of this test indicated variations between teachers significant at the .05 level for four of the six groups of observers. All F values for these four groups were significant at the .01 level except one; this was level one of the untrained secondary science teacher group. All F values for trained observers were significant at the .01 level. The only non-trained observers with significant F values was the secondary school science teacher observer team.

Reliability coefficients were computed for all significant F value observer groups. These values are given on each of the analysis of variance Tables III through XX. A composite of all reliability co-efficients is shown in Table XXI. The highest reliability was .6986 for the trained elementary teachers on level two of the taxonomy. The lowest reliability coefficient, of those with significant F values, was the untrained secondary school science teachers with a level one value of .2122.

The coefficients of observer agreements were computed by using the Scott formula for each of the observer teams and each level of the taxonomy (50). These coefficients are presented with the reliability coefficients in Tables III through XX. A composite of all coefficients of observer agreements is shown in Table XXII. The highest coefficient of observer agreement was .7777 recorded by the trained secondary science teacher group. The greatest combination of estimated reliability index and coefficient of observer agreement was recorded by the trained elementary teachers on level two of the taxonomy.

Descriptive Analysis of the Data

The ability of observers to agree on the classification of the filmed scenes is shown to increase as the number of divisions of the taxonomy decrease even though the probability of agreement by chance is removed.

The data indicate that level two is the most reliable of the three levels. Each of the four groups of observers with significant F values scored highest on this level.

Summary

Statistical calculations using analysis of variance techniques for reliability estimation indicate significant differences in classroom activity scenes as observed and recorded by four of the six observer groups. The trained elementary teachers make the best observers as measured by the reliability coefficients and observer agreement statistics.

The data indicate that significant differences in the filmed classroom scenes were attributed to teacher differences by four of the six observer groups using the taxonomy of classroom activity.

TABLE III

ANALYSIS OF VARIANCE FOR TRAINED HOUSEWIFE OBSERVERS, LEVEL I

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	Calculated F Ratio	Tabular F Ratio
Total	59	62,555.6000			
Among Means of Different Observed Classrooms	9	46,817.9333	5,201.9925		
Between Observers	5	2,179.6000	435.9200		
Observers X Classrooms	45	13,558.0667	301.2903	7.0563	

Reliability Coefficient = .5023

Coefficient of Observer Agreement = .4432

TABLE IV

ANALYSIS OF VARIANCE FOR TRAINED HOUSEWIFE OBSERVERS, LEVEL II

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	Calculated F Ratio	Tabular F Ratio
Total	59	7,646.1834			
Among Means of Different Observed Classrooms	9	5,897.0167	655.2240		
Between Observers	5	221.6834	44.3366		
Observers X Classrooms	45	1,527.4833	33.9440	8.3701	

Reliability Coefficient = .5512

Coefficient of Observer Agreement = .5710

TABLE V

ANALYSIS OF VARIANCE FOR TRAINED HOUSEWIFE OBSERVERS, LEVEL III

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	Calculated F Ratio	Tabular F Ratio
Total	59	812.8500			
Among Means of Different Observed Classrooms	9	566.0166	62.8907		
Between Observers	5	36.1500	7.2300		
Observers X Classrooms	45	210.6834	4.6818	5.2796	

Reliability Coefficient = .4164

Coefficient of Observer Agreement = .7036

TABLE VI

ANALYSIS OF VARIANCE FOR TRAINED ELEMENTARY TEACHER OBSERVERS, LEVEL I

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	Calculated F Ratio	Tabular F Ratio
Total	59	66,434.9834			
Among Means of Different Observed Classrooms	9	51,271.1500	5,696.7944		
Between Observers	5	764.7039	152.9407		
Observers X Classrooms	45	14,399.1295	319.9806	12.0459	

Reliability Coefficient = .6480

Coefficient of Observer Agreement = .6521

TABLE VII

ANALYSIS OF VARIANCE FOR TRAINED ELEMENTARY TEACHER OBSERVERS, LEVEL II

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	Calculated F Ratio	Tabular F Ratio
Total	59	8,373.2500			
Among Means of Different Observed Classrooms	9	6,532.4166	725.8240		
Between Observers	5	43.7500	8.7500		
Observers X Classrooms	45	1,797.0833	39.9351	14.9085	

Reliability Coefficient = .6986

Coefficient of Observer Agreement = .7025

TABLE VIII

ANALYSIS OF VARIANCE FOR TRAINED ELEMENTARY TEACHER OBSERVERS, LEVEL III

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	Calculated F Ratio	Tabular F Ratio
Total	59	768.9334			
Among Means of Different Observed Classrooms	9	483.9334	53.7704		
Between Observers	5	11.7334	2.3466		
Observers X Classrooms	45	273.2666	6.0725	6.3867	

Reliability Coefficient = .4730

Coefficient of Observer Agreement = .7110

TABLE IX

ANALYSIS OF VARIANCE FOR TRAINED SECONDARY SCIENCE TEACHER OBSERVERS, LEVEL I

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	Calculated F Ratio	Tabular F Ratio
Total	59	78,079.2500			
Among Means of Different Observed Classrooms	9	45,077.7500	5,008.6389		
Between Observers	5	3,481.9500	696.3900		
Observers X Classrooms	45	29,519.5500	655.99	3.7035	

Reliability Coefficient = .3106

Coefficient of Observer Agreement = .6402

TABLE X

ANALYSIS OF VARIANCE FOR TRAINED SECONDARY SCIENCE TEACHER OBSERVERS, LEVEL II

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	Calculated F Ratio	Tabular F Ratio
Total	59	8,934.8500			
Among Means of Different Observed Classrooms	9	5,674.6833	630.2036		
Between Observers	5	351.7500	70.3500		
Observers X Classrooms	45	2,908.4167	64.6314	4.6688	

Reliability Coefficient = .3794

Coefficient of Observer Agreement = .6848

TABLE XI

ANALYSIS OF VARIANCE FOR TRAINED SECONDARY SCIENCE TEACHER OBSERVERS, LEVEL III

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	Calculated F Ratio	Tabular F Ratio
Total	59	868.9333			
Among Means of Different Observed Classrooms	9	520.6000	57.8444		
Between Observers	5	40.2000	8.0400		
Observers X Classrooms	45	308.1333	6.8474	3.8854	

Reliability Coefficient = .3247

Coefficient of Observer Agreement = .7777

TABLE XII

ANALYSIS OF VARIANCE FOR UNTRAINED HOUSEWIFE OBSERVERS, LEVEL I

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	Calculated F Ratio	Tabular F Ratio
Total	59	71,637.333			
Among Means of Different Observed Classrooms	9	21,680.3333	2,408.9259		
Between Observers	5	9,452.1333	1,890.4266		
Observers X Classrooms	45	40,504.8667	900.1081	.8632	

Reliability Coefficient = Not Significantly Different From Zero

Coefficient of Observer Agreement = .3804

TABLE XIII

ANALYSIS OF VARIANCE FOR UNTRAINED HOUSEWIFE OBSERVERS, LEVEL II

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	Calculated F Ratio	Tabular F Ratio
Total	59	8,346.7333			
Among Means of Different Observed Classrooms	9	2,773.7333	308.1926		
Between Observers	5	949.7333	189.9466		
Observers X Classrooms	45	4,623.2667	102.7392	1.0529	

Reliability Coefficient = Not Significantly Different From Zero

Coefficient of Observer Agreement = .4302

TABLE XIV

ANALYSIS OF VARIANCE FOR UNTRAINED HOUSEWIFE OBSERVERS, LEVEL III

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	Calculated F Ratio	Tabular F Ratio
Total	59	902.7333			
Among Means of Different Observed Classrooms	9	334.4000	37.1556		
Between Observers	5	93.3333	18.6666		
Observers X Classrooms	45	475.0000	10.5555	1.2714	

Reliability Coefficient = Not Significantly Different From Zero

Coefficient of Observer Agreement = .5554

TABLE XV

ANALYSIS OF VARIANCE FOR UNTRAINED ELEMENTARY TEACHER OBSERVERS, LEVEL I

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	Calculated F Ratio	Tabular F Ratio
Total	59	60,126.9833			
Among Means of Different Observed Classrooms	9	24,817.8166	2,757.5352		
Between Observers	5	10,522.6833	2,104.5366		
Observers X Classrooms	45	24,786.4834	550.8107	1.0384	

Reliability Coefficient = Not Significantly Different From Zero

Coefficient of Observer Agreement = .3861

TABLE XVI

ANALYSIS OF VARIANCE FOR UNTRAINED ELEMENTARY TEACHER OBSERVERS, LEVEL II

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	Calculated F Ratio	Tabular F Ratio
Total	59	7,217.6500			
Among Means of Different Observed Classrooms	9	3,139.4833	348.8315		
Between Observers	5	1,159.5500	231.9100		
Observers X Classrooms	45	2,916.6167	64.8137	1.1756	

Reliability Coefficient = Not Significantly Different From Zero

Coefficient of Observer Agreement = .4726

TABLE XVII

ANALYSIS OF VARIANCE FOR UNTRAINED ELEMENTARY TEACHER OBSERVERS, LEVEL III

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	Calculated F Ratio	Tabular F Ratio
Total	59	750.5833			
Among Means of Different Observed Classrooms	9	334.7500	37.1944		
Between Observers	5	108.6833	21.7366		
Observers X Classrooms	45	307.1500	6.8255	1.3022	

Reliability Coefficient = Not Significantly Different From Zero

Coefficient of Observer Agreement = .5777

TABLE XVIII

ANALYSIS OF VARIANCE FOR UNTRAINED SECONDARY SCIENCE TEACHER OBSERVERS, LEVEL I

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	Calculated F Ratio	Tabular F Ratio
Total	59	69,496.7334			
Among Means of Different Observed Classrooms	9	41,699.4000	4,633.2667		
Between Observers	5	6,483.9334	1,296.7866		
Observers X Classrooms	45	21,313.4000	473.6311	2.6170	

Reliability Coefficient = .2122

Coefficient of Observer Agreement = .4232

TABLE XIX

ANALYSIS OF VARIANCE FOR UNTRAINED SECONDARY SCIENCE TEACHER OBSERVERS, LEVEL II

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	Calculated F Ratio	Tabular F Ratio
Total	59	8,310.9834			
Among Means of Different Observed Classrooms	9	5,121.8167	569.0907		
Between Observers	5	705.1534	141.0306		
Observers X Classrooms	45	2,484.0133	55.2002	2.9001	

Reliability Coefficient = .2405

Coefficient of Observer Agreement = .5484

TABLE XX

ANALYSIS OF VARIANCE FOR UNTRAINED SECONDARY SCIENCE TEACHER OBSERVERS, LEVEL III

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Squares	Calculated F Ratio	Tabular F Ratio
Total	59	792.0000			
Among Means of Different Observed Classrooms	9	502.0000	55.7778		
Between Observers	5	73.0000	14.6000		
Observers X Classrooms	45	217.0000	4.8222	2.8718	

Reliability Coefficient = .2377

Coefficient of Observer Agreement = .7296

TABLE XXI

COMPOSITE RELIABILITY COEFFICIENTS

Trained Housewives (Observers 1-6)	
Level I (37 classes)	.5023
Level II (12 classes)	.5512
Level III (4 classes)	.4164
Trained Elementary Teachers (Observers 7-12)	
Level I (37 classes)	.6480
Level II (12 classes)	.6986
Level III (4 classes)	.4730
Trained Secondary Science Teachers (Observers 13-18)	
Level I (37 classes)	.3106
Level II (12 classes)	.3794
Level III (4 classes)	.3247
Untrained Secondary Science Teachers (Observers 31-36)	
Level I (37 classes)	.2122
Level II (12 classes)	.2405
Level III (4 classes)	.2377

TABLE XXII

COEFFICIENTS OF OBSERVER AGREEMENT

Trained Housewives (Observers 1-6)	
Level I	.4432
Level II	.5710
Level III	.7036
Trained Elementary Teachers (Observers 7-12)	
Level I	.6521
Level II	.7025
Level III	.7110
Trained Secondary Science Teachers (Observers 13-18)	
Level I	.6402
Level II	.6848
Level III	.7777
Untrained Housewives (Observers 19-24)	
Level I	.3804
Level II	.4302
Level III	.5554
Untrained Elementary Teachers (Observers 25-30)	
Level I	.3861
Level II	.4726
Level III	.5777
Untrained Secondary Science Teachers (Observers 31-36)	
Level I	.4232
Level II	.5484
Level III	.7296

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Summary

The purpose of this study was to design and test an observational instrument for use in the collection of data pertaining to the classroom activities of junior high school science students. The instrument, Taxonomy of Classroom Activity, was designed for recording short time duration observations made by observers located outside of the classroom. The taxonomy was developed by revising the twenty-three item Variety Section of the Code Digest into a three level taxonomy.

The instrument was evaluated by showing sixty filmed classroom scenes, taken from the hallway, to thirty-six observers. The filmed scenes were taken from ten teachers' classrooms on six different occasions. The observers were divided into two groups and three teams within each group. One of the groups had no formal training or practice in the use of the instrument but were provided with the manual in Appendix A. The second observer group received formal training and practice in addition to the manual. Each observer group consisted of six secondary science teachers, six elementary teachers, and six housewives. Reliability indexes and coefficients of observer agreement were computed for each observer team by use of analysis of variance and the Scott formula.

Conclusions

On the basis of experimental evidence there is reason to reject the following research hypothesis:

1. The previous training and background of observers do not alter the reliability of the observations using the taxonomy of classroom activity.

The data indicate that the non-science observers obtained more reliable scores than the science teachers.

2. The training of observers in the use of the taxonomy of classroom activity by formal instruction does not alter the reliability of the observations.

The untrained science teachers were the only untrained observers with reliabilities significantly different from zero.

3. The training of observers in the use of the taxonomy of classroom activity by formal instruction does not alter the coefficients of observer agreement.

The data for the trained observers contained seven of nine observer agreement coefficients in excess of 60%, while the non-trained group contained only one score in excess of 60%.

There appears to be no evidence to support a rejection of research hypothesis two: The previous training and background of observers do not alter the coefficient of observer agreement. The data indicate, in a descriptive way, a closer relationship to training in the use of the instrument than to background and previous training. The importance of this instrument training is rather striking for all groups except the secondary science teachers. The reasons for the low scores by the secondary science teachers as compared to the elementary teachers and the housewives is not apparent from the data.

Observers without formal training sessions may have been limited by the manual and a revision of the document could possibly improve the reliability of data collected by untrained observers.

The study addressed itself to the following questions:

1. How reliabile is an observation made from the hall?

The data indicate that reliable information, of the nature for which this taxonomy was designed, may be obtained by observing from the hall.

2. How valid is such an observation using the proposed taxonomy?

The validity of this instrument, as pointed out by Medley and Mitzel (1), depends upon the inferences and conclusions drawn from data collected in this way. The taxonomy of classroom activity was designed to find out, by observational methods, what kinds of activities take place in junior high school science classes. The instrument makes no attempt to judge "good" or "poor" activities or evaluate teachers in the sense of teaching effectiveness.

3. Can observers agree on the classification of a scene viewed from the hall?

The data indicate several agreements in excess of 70%.

4. Is formal training in the use of the instrument helpful in increasing the reliability and coefficients of observer agreement?

The data indicate differences in favor of the trained observers in all cases.

5. Is formal college training in education necessary for observers?

The data seem to indicate slightly higher scores for the trained elementary teachers as compared to the trained housewives, but the difference may be due to something other than formal college training. The evidence seems to indicate that high school graduate housewives, with proper training in use of the instrument, may do quite satisfactory work as observers.

6. Is formal college training in science and education essential for maximum reliability, and observer agreement?

Since the highest scores were obtained by elementary teachers and housewives, it would appear that formal training in science is not essential for observers.

Recommendations

There were six developmental phases in this project described in Chapter III. The following recommendations are made pertaining to each phase:

1. Development of the taxonomy

The taxonomy should be revised to include a category for class discussion. This may be added to the four basic divisions or added as a second level division under listening and viewing.

2. The Observer Training Manual

The training manual in its present form is inadequate as a single training device. The results of this study indicate very definite need for training in excess of this manual.

3. Selection and filming of classroom scenes

Many technical problems were encountered in the sound synchronization using the 16 millimeter system. Future researchers may wish to consider video tape systems.

4. Observer selection and training

The evidence indicates that users of this instrument must be trained in formal sessions for significant results. Observers would probably do better than those in this study if additional training time was required. The optimum training period is not yet known for this instrument.

5. Collection of the data

The present system of recording is awkward for observers moving around in hallways. A more compact way of recording is needed.

6. Analysis of data

In situations where every observer does not observe the same scenes, it will be necessary to include the variance due to teacher visit differences.

Finally, the investigator would like to see the use of this instrument on a wide geographical area basis to determine the types and frequencies of junior high school science activities.

A SELECTED BIBLIOGRAPHY

1. Medley, Donald M., and Harold E. Mitzel. "Measuring Classroom Behavior by Systematic Observation," Chapter 6, Handbook of Research on Teaching, American Education Research Association, Rand McNally and Co. 1963.

2. Flanders, N. A. "Teacher Influence, Pupil Attitudes and Achievement," Studies in Interaction Analysis. Final Report, Cooperative Research Project No. 397, U.S. Office of Education, 1960.

3. Wright, Muriel J. "Development of an Instrument for Studying Verbal Behaviors in a Secondary School Mathematics Classroom." Journal of Experimental Education, No. 28, (1959) 103.

4. Sears, Pauline. "The Effect of Classroom Conditions on the Strength of Self-Esteem, Achievement Motivation and Work Output of Elementary School Children." Final Report, Cooperative Research Project No. 873, U.S. Office of Education, 1962.

5. Simon, Anita. "Research and Training Reviews," Classroom Interaction Newsletter, No. 1 (1965) 7.

6. Fuller, Frances. "The Bales System and Flanders System for 8 Millimeter Films," Interaction Newsletter No. 1 (1965) 14.

7. Michel, M. "Maywood's Mobile T.V. Unit," Audio Visual Instructor, Vol. 11 (January, 1966) 27.

8. Cornell, F. G. et al., "An Exploratory Measurement of Individualities of Schools and Classrooms," (Urbana, Ill. Bureau of Educational Research, College of Education, University of Illinois, 1952).

9. Watson, F. G. "Research on Teaching Science," Chapter 20, Handbook of Research on Teaching, American Education Research Association, Rand McNally and Co., 1963, 260.

10. La Shier, William S., Jr. "An Analysis of Certain Aspects of the Verbal Behavior of Student Teachers of Eighth Grade Students Participating in a BSCS Laboratory Block." Unpublished doctoral dissertation, The University of Texas, 1965.

11. Molchen, Kenneth J. "A Study in the Relationship Between the Objectives and the Actual Classroom Behavior in Science," Interaction Newsletter, No. 1, (1965) 21.

12. Mathews, Charles C. "The Identification of Non-Random Changes in Verbal Behaviors of Science Student Teachers as Related to the Verbal Behaviors of the Associated Cooperating Teachers," Interaction Newsletter, No. 1, (1965) 19.

13. Leake, John B. "A Study of Attitudes of Elementary Teachers Toward Science." Unpublished doctoral dissertation, Oklahoma State University (1966).

14. Bloom, B. S., and Krathwohl. "Taxonomy of Educational Objectives: Classification of Educational Goals," Handbook I, The Cognitive Domain, Longman, Green, 1956.

15. Openshaw, Karl M., and Frederick R. Cyphert. "Development of a Taxonomy for the Classification of Teacher Classroom Behavior." Final Report, Cooperative Reserach Project No. 2288, U.S. Office of Education, (1966).

16. Horn, Earnest. "Distribution of Opportunity for Participation Among Various Pupils in Classroom Recitation." Contributions to Education, No. 67, New York: Teachers College, Columbia University, 1914.

17. Puckett, R. C. "Making Supervision Objective." School Review, Vol. 36, (1938).

18. Wrightstone, J. W. "Appraisal of Newer Elementary School Practices." New York: Teachers College, Columbia University, (1938).

19. Jersild, Arthur T., and Margaret F. Meigs. "Direct Observation as a Research Method." Review of Educational Research, Vol. 9, (December, 1939) 472-82.

20. Carlson, J. S. "Sex Differences in Conversation." Journal of Applied Psychology, Vol. 20, (December, 1936) 727-35.

21. Landis, Carney. "National Differences in Conversation." Journal of Abnormal and Social Psychology, Vol. 21, (January, 1927) 354-57.

22. Landis M. H., and H. E. Burtt. "The Study of Conversations," Journal of Comparative Psychology, Vol. 4, (1924) 81.

23. Arrington, Ruth E. "Some Technical Aspects of Observer Reliability as Indicated in Studies of the 'Talkies.'" American Journal of Sociology, Vol. 38, (November, 1932) 409-17.

24. Olson, W. C., and E. M. Cunningham. "Time Sampling Techniques." Child Development, Vol. 5, (March, 1934) 41-58.

25. Thorndike, Robert L., John J. Loftus, and Bernard Goldman. "Observation of the Behavior of Children in Activity and Control Schools." Journal of Experimental Education, Vol. 10, (December, 1941) 138-145.

26. Barr, Arvil S. "Characteristic Differences in the Teaching Performance of Good and Poor Teachers of Social Studies. Illinois: Public School Publishing Company, 1929.

27. Jayne, C. D. "A Study of the Relationship Between Teaching Procedures and Educational Outcomes." Journal of Experimental Education, Vol. 14, (1945) 101-134.

28. Morsh, J. E. "Development Report - Systematic Observation of Instructor Behavior." USAF Pers. Train. Res. Cent. Develpm. Rept., AFPTRC-TN 56-52, 1956.

29. Ryan, David G. "Some Relationships Between Pupil Behavior and Certain Teacher Characteristics." Journal of Educational Psychology, Vol. 52, (1961) 82-90.

30. ________. "Prediction of Teacher Effectiveness." Encyclopedia of Educational Research, Ed. Chester W. Harris. New York: McMillian, 1960, 1486-91..

31. Anderson, H. M. "A Study of Certain Criteria of Teaching Effectiveness." Journal of Experimental Education, Vol. 23, (1954) 67.

32. Jones, R. D. "The Prediction of Teaching Efficiency from Objective Measures." Journal of Experimental Education, Vol. 15, (1946) 98.

33. Domas, S. J., and D. V. Tiedman. "An Annotated Bibliography." Journal of Experimental Education, Vol. 19, (December, 1950) 99-218.

34. Barr, A. S. "Teacher Effectiveness: A Summary of Investigations." Madison, Wisconsin: Dembar Publications, Inc., 1961.

35. Getzels, J. W., and P. W. Jackson. "The Teacher's Personality and Characteristics." Handbook of Research on Teaching, American Educational Research Association. Rand McNally and Co., (1963).

36. Thomas, D. S. et al., "Some New Techniques for Studying Social Behavior." Child Development Monographs, No. 1, New York: Teachers College, Columbia University, 1929. 203.

37. Anderson, H. H., and Helen M. Brewer. Studies of Teachers' Classroom Personalities. I. "Dominative and Socially Integrative Behavior of Kindergarten Teachers." _Applied Psychology Monograph_, No. 6, (1945).

38. Anderson, H. H., and J. E. Brewer. "Studies of Teachers' Dominative and Inegrative Contacts on Children's Classroom Behavior." _Applied Psychology Monograph_, No. 8, (1946).

39. Anderson, H. H., J. E. Brewer, and Mary F. Reed. Studies of Teachers' Classroom Personalities. III. Followup Studies of the Effects of Dominative and Integrative Contacts on Children's Behavior." _Applied Psychology Monograph_ No. 11, (1946).

40. Withall, John. "The Development of A Technique for the Measurement of Social Emotional Climate in Classrooms." _Journal of Experimental Education_, Vol. 17, (1949), 347-61.

41. ________. "The Development of a Climate Index." _Journal of Educational Research_, Vol. 45, (1951) 93-99.

42. Anderson, Harold. "Dominative and Socially Integrated Behaviors." _Child Behavior and Development_. New York: McGraw Hill, 1943.

43. Hughes, M. _Development of the Means for the Assessment of the Quality of Teaching in Elementary Schools_. Salt Lake City: University of Utah Press, (1949).

44. Parakh, Jal S. "A Study of Teacher-Pupil Interaction in High School Biology Classes." A paper presented at the annual meeting of the National Association for Research in Science Teaching, February, 1966, Chicago, Illinois. The research reported herein was supported by the Cooperative Research Program of the Office of Education, U.S. Department of Health, Education, and Welfare (Cooperative Research Project No. S-269).

45. Medley, Donald, and Harold Mitzel. "A Technique for Measuring Classroom Behavior." _Journal of Educational Psychology_, Vol. 49, (1958) 86-92.

46. Travers, R. M. V. et al., "Measured Needs of Teachers and Their Behavior in the Classroom." Final Report, Cooperative Research Project No. 444, U.S. Office of Education, 1961.

47. Wandt, Edwin, and Leonard Ostreicher. "Validity of Samples of Classroom Behavior for the Measurement of 'Social-Emotional Climate.'" _Psychological Monographs_, Vol. 68, (1954).

48. Medley, Donald, and Harold Mitzel. "Application of Analysis of Variance to the Estimation of the Reliability of Observation of Teachers' Classroom Behavior." _Journal of Experimental_ Education, Vol. 27, (September, 1958) 23-35.

49. Perkins, Hugh. "A Procedure for Assessing the Classroom Behavior of Students and Teachers." American Educational Research Journal, Vol. 1, (November, 1964), 249-260.

50. Scott, W. A. "Reliability of Content Analysis: The Case of Nominal Scale Coding." Public Opinion Quarterly, Vol. 19, (1955) 321-325.

51. Ostle, Bernard. Statistics in Research. Ames: Iowa State University Press, 1963.

52. Jackson, R. W. B., and G. A. Ferguson. Studies on the Reliability of Tests, Bulletin 12, (Toronto: Department of Educational Research, University of Toronto, 1941) 132.

53. Hoyt, Cyril. "Test Reliability Estimated by Analysis of Variance." Psychometrika, Vol. 6, (1941) 153-160.

54. Alexander, H. W. "The Estimation of Reliability When Several Trials Are Available." Psychometrika, Vol. 12, (1947) 79-99.

APPENDIX A

OBSERVERS MANUAL FOR THE TAXONOMY OF CLASSROOM ACTIVITY

OBSERVERS MANUAL

for

A TAXONOMY OF CLASSROOM ACTIVITY

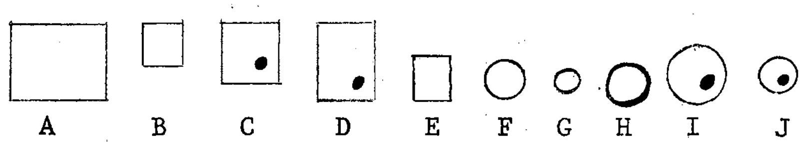
by

Neal Jay Holmes

A Taxonomy of Classroom Activity

The taxonomy of classroom activity is an instrument under experimental study for possible use in certain types of educational research. The objective is to classify the activity in a classroom at one precise time without the interaction of the observer upon the classroom being observed. This requirement makes it essential for the observer to be located in the hall or other outside observation post. The taxonomy is a classification system designed specifically to determine, without classroom interaction by observers, the type and frequency of classroom activities by different classes in various learning situations.

Very simply we want to place the various activities into categories much as one would do in sorting beads or other objects. One example of such a taxonomy could be constructed using the following types of objects:



A glance through this collection of objects reveals two basic shapes (round and square). This would provide us with the initial two categories for dividing the collection. The first division of our taxonomy, noted 1.00 and 2.00, would be round (1.00) and square (2.00). This does not discriminate very well; in other words this doesn't help us to sort out a particular object, and further sorting appears to be necessary. Another look at the objects reveals that some of them have holes while others do not. Using the property of holes and no-holes, we can sort the circular and square objects into two additional categories. We would now have the following divisions:

Square - Without Holes. . . . . Noted 1.10

Square - With Holes . . . . Noted 1.20

Round - Without Holes . . . . Noted 2.10

Round - With Holes. . . . . Noted 2.20

The taxonomy is often most convenient to use in the numeral notation form since economy of space in complex classes is desirable. Let us see how this would work with our system as it is developed thus far.

1.00 All Square Objects (Includes A, B, C, D, E)

- 1.10 All Squares without holes (A, B, E)
- 1.20 All Squares with holes (C, D)

2.00 All Circular Objects (Including F, G, H, I, J)

- 2.10 All Circular Objects without holes (F, G, H)
- 2.20 All Circular Objects with holes (I, J)

How well does this sort our original collection? The answer to this question depends on how precisely we would want to identify the object or group of objects. If we knew the color of these objects, for example, they could then be divided using this characteristic.

For example: Suppose (A) red, (B) blue (C) green (D) red (E) green (F) red (G) green (H) blue (I) red (J) blue. We may now divide our group a third time sorting by colors.

The taxonomy now has 12 divisions based on three basic characteristics: shape, structure, and color.

The complete taxonomy consists of the following categories:

1.00 Square Objects

- 1.10 Square Objects without holes

1.11 Red Square Objects without holes

1.12 Blue Square Objects without holes

1.13 Green Square Objects without holes

1.20 Square Objects with holes

1.21 Red Square Objects with holes

1.22 Blue Square Objects with holes

1.23 Green Square Objects with holes

2.00 Circular Objects

2.10 Circular Objects without holes

2.11 Red Circular Objects without holes

2.12 Blue Circular Objects without holes

2.13 Green Circular Objects without holes

2.20 Circular Objects with holes

2.21 Red Circular Objects with holes

2.22 Blue Circular Objects with holes

2.23 Green Circular Objects with holes

Using this system it is now possible to classify the items (A through J) without any two items having the same key numerals. From the information on pages one and two we can arrive at the following classification numerals for the objects: A-1.11, B-1.12, C-1.23, D-1.21, E-1.13, F-2.11, G-2.13, H-2.12, I-2.21, J-2.22. This is an idealized taxonomy for our collection of objects, it will be apparent that all material objects will not fit this system and many revisions would be necessary if one desired to classify a different group of objects. It should also be apparent that other equally valid systems of classification could be used with the above items. The objective for this study is using the taxonomy not designing one.

Using the taxonomy outlined on pages one, two, and three, classify the following items by recording the appropriate code numerals in the spaces provided.

(A) A red square with a hole . . . . ______________

(B) A green square without a hole. . . . . ______________

(C) A blue circular object without a hole. . . . . ______________

(D) A red circular object with a hole. . . . . . . ______________

(E) A blue square object with a hole . . . . ______________

(F) A green circular object with a hole. . . . . . ______________

(See page 19 for the correct answers.)

Review any incorrect responses to the above questions before continuing in this manual.

THE CLASSROOM ACTIVITY TAXONOMY

Classroom activity involves the complexities encountered in dealing with human behavior and for that reason only a limited part of the total behavior is considered in this taxonomy. This taxonomy is concerned with the variety of classroom activity, including: (1) what type of things students do in the classroom, and (2) the relative frequency of each activity. The system is not designed to evaluate "good" activities or determine good teaching.

The Classroom Activity Taxonomy is constructed similarly to the one just described. The major activities are:

Listening and Viewing . . . . .Noted 1.00

Reading and Writing . . . . .Noted 2.00

Manipulative Work . . . . .Noted 3.00

Diverse Activities. . . . . .Noted 4.00

Each major activity is subdivided into three divisions as shown below:

1.00 Students listening and/or viewing

1.10 Students listening

1.20 Students viewing

1.30 Mixed student listening and viewing

2.00 Students reading and/or writing

2.10 Students reading

2.20 Students writing

2.30 Students reading and writing

3.00 Students performing manipulative tasks

3.10 Students working in groups of three or more

3.20 Students working in groups of two

3.30 Students working individually

4.00 Students performing two or more of the above tasks

4.10 Students working in groups of two or more

4.20 Students working individually

4.30 Students working individually and in groups

The observer must make a judgment of the observed behavior and select one of the four major categories. The remaining three categories are eliminated for that observation. After a major category has been chosen, the observer must select one of the three subdivisions. The third level of separation involves from two to four possibilities depending on the major category. The complete taxonomy is shown below:

1.00 Listening and/or viewing

1.10 Listening

1.11 To a mechanical device

1.12 To a teacher

1.13 To a student

1.20 Viewing

1.21 A physical object or reproduction

1.22 A teacher demonstration

1.23 A student demonstration

1.30 Mixed listening and viewing

1.31 Audio-Visual apparatus

1.32 Listening to teacher viewing other

1.33 Listening to student viewing other

2.00 Reading and Writing

2.10 Reading

2.11 Reading textbooks

2.12 Non-textbooks

2.20 Writing

2.21 Manual or workbook

2.22 Student paper or notebook

2.23 Testing (essay)

2.30 Mixed reading and writing

2.31 Text, notebook and/or workbook

2.32 Objective testing

2.33 Other reading and writing

3.00 Students performing manipulative tasks

3.10 Students working in groups of three or more

3.11 Using similar materials

3.12 Using different materials

3.13 Class out of classroom

3.20 Students working in groups of two

3.21 Using similar materials

3.22 Using different materials

3.23 Unclassified

3.30 Students working individually

3.31 Using similar materials

3.32 Using different materials

4.00 Students performing two or more of the above tasks

4.10 Students working in groups of two or more

4.11 Activities involving 1.00, 2.00, and 3.00

4.12 Activities involving 1.00 and 2.00

4.13 Activities involving 1.00 and 3.00

4.14 Activities involving 2.00 and 3.00

4.20 Students working individually

4.21 Activities involving 1.00, 2.00, and 3.00

4.22 Activities involving 1.00 and 2.00

4.23 Activities involving 1.00 and 3.00

4.24 Activities involving 2.00 and 3.00

4.30 Students working individually and in groups

4.31 Activities involving 1.00, 2.00, and 3.00

4.32 Activities involving 1.00 and 2.00

4.33 Activities involving 1.00 and 3.00

4.34 Activities involving 2.00 and 3.00

DEFINITION OF TERMS AND EXAMPLES

1.00 Listening and/or viewing activities are usually identifiable by the characteristics of the sounds heard in the hallway. The position of the students' heads as they view the object or person is also helpful in classification. The film strip or slide projection activity (classified 1.21) may be silent, while a teacher giving a lecture or discussion from a chalkboard diagram (classified 1.32) may be quite obvious from the sound of the teacher's voice. A student talking in a discussion session would involve primarily listening (1.13) but situations may be observed where obvious viewing is integrated with the listening; in these cases it would be necessary to classify the activity 1.33. The observer must remember that this system is a short time duration observation and the behavior tone is the one to classify.

2.00 Reading and writing activities require a visual contact with the student and his or her work area. The reading of identical books would be classified as reading textbooks (2.11). If the students are reading materials of various types such as paper backed books, hard backed books, and loose paper materials, the classification would be 2.12. The only division of the writing activities is the distinction between text and workbook combination (2.21), notebook and other copy writing (2.22), and essay test or report writing where no indication is given of copying from source materials (2.23).

Situations involving both reading and writing would be exemplified by the presence of book materials as well as writing materials. One such example would be workbook writing requiring periodic reference to textbooks (2.31). Another example would be a situation involving multiple type test questions requiring reading and writing closely integrated (2.32). Reading and writing activities not fitted to the above should be classified 2.33.

3.00 Students performing manipulative tasks. This category is of particular interest to disciplines involving some elements of laboratory tasks such as shop, art, music, and science. All classes viewed by observers in this study will be from the sciences taught in the junior high school. The manipulative category is defined as student work with physical materials other than pencil and paper. The first subdivision within the manipulative area involves the description of the type of grouping of students during the activity. There may be occasions when more than one size of group is represented; in that event, the code number representing the greatest group is to be used. In cases where only one group is observed it will be assumed that others are grouped similarly. The second subdivision is designed to distinguish students or student groups and the type of materials they are working with. The term similar in this particular case means materials for the same kind of activity, such as microscopes. The microscopes may be of different manufacture, etc. but the objective is the same type of activity. Two additional classes are inserted in this major category to accommodate the field trip (3.13), and the unclassifiable observation (3.23). Category 3.23 should be used with care and every effort should be made to be certain no category has been overlooked before deciding the behavior is unclassifiable.

4.00 Diversity: This category is designed to classify those situations involving mixed or diverse activities. For example, if some students are watching the teacher while others are working with laboratory activities, the correct classification is 4.13. The activity is mixed between viewing and manipulating with both being performed in a group situation. A reading group and a laboratory group would be classified 4.14. The 4.20 series is used only when everyone is working singly. The 4.30 series is used when there are groups and individuals working on various situations, such as reading, writing, and laboratory work 4.34, or listening and laboratory work 4.33.

Classify the following scenes and refer to the discussion of these pictures on page 16 for possible coding. Since no sound is available, some situations may be coded in more than one way.

Picture No. 1

This picture illustrates a situation often observed in classrooms with the teacher pointing out a visual reference. Since no sound is available, we do not know if listening is part of this situation or not. The scene must, very definitely, be classified in the 1.00 section, and it appears obvious that the students are viewing. If the teacher is talking, the classification would be 1.32. If the teacher is not talking, then we would classify the scene 1.22.

Picture No. 2

This is a rather clear example involving a writing situation and would have to be in the 2.00 area. Each student is writing, therefore, we must use the 2.20 category and since they are writing on notepaper, we would classify this scene as 2.22.

Picture No. 3

This situation involves either a student demonstration or an experiment in progress. Three students are working with physical materials; the table appears to be a teacher's laboratory demonstrating table, and the student's head in the lower part of the picture implies that students are watching. The author classifies the scene 1.23. You may have classified the scene 3.11. Usually you will be able to see more than one station from the hallway and would not have to infer what other students are doing in the classroom.

Picture No. 4

This appears to be a laboratory group working in fours. There appears to be manipulation of apparatus in groups of three or more which would fit the classification 3.10. You may notice the pen in hand of one of the students and decide to classify this activity 4.14 involving manipulation and writing. This does not appear to be writing but more of a record taking action as part of the manipulation. We do not know what other groups are doing since only the one is visible and would have to infer the same type of activity and classify the picture 3.11.

Picture No. 5

This is a viewing situation and probably involves listening. The area is 1.30 but without the sound it would be impossible to tell the source. It would not be incorrect to classify this scene 1.10 if you assumed there was no sound.

Picture No. 6

A textbook reading situation classified 2.11.

Picture No. 7

Teacher talking students viewing and listening, classified 1.32. It could be classified 1.22 if no-one is talking or 1.33 if a student is talking.

Picture No. 8

Appears to be strictly viewing but sound may be coming from somewhere or someone. Observers should agree that this is a 1.00 activity. The sound should tell us if it is a 1.20 or a 1.30 situation. Since all of the children appear to be looking in the same direction, viewing would appear to be a correct choice. The listening would have to come from the sound source.

Picture No. 9

The class appears to be viewing the photographer at the moment but the work underway by the materials out on the desks would infer the reading and writing situation with textbooks, notebooks, and workbooks 2.31.

Picture No. 10

Individual manipulative tasks, using similar materials, classification 3.31.

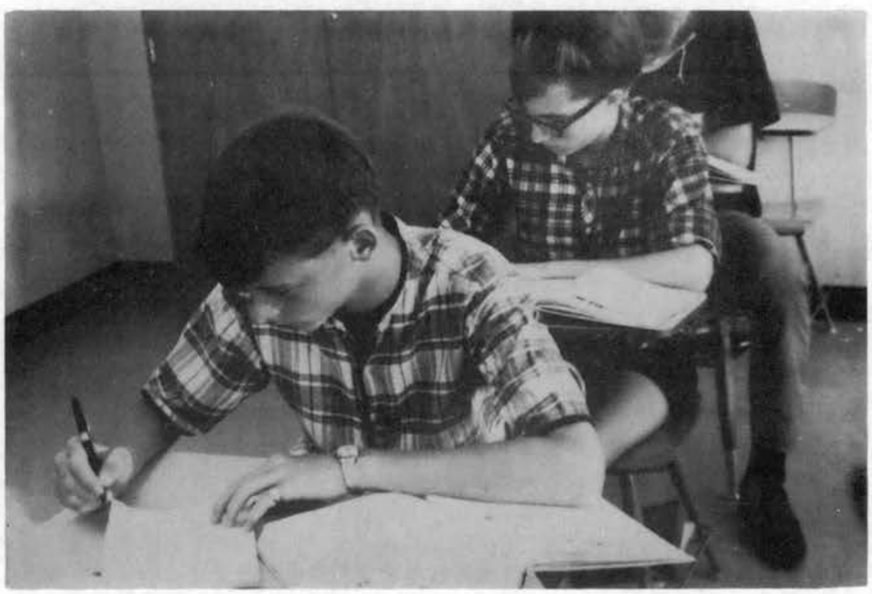
The pictures are provided to familiarize observers with a few of the various situations found in the classrooms. You will be viewing 16 millimeter sound films taken from the hallway and providing many more clues to the exact nature of the activity than can be provided with still pictures.

If you did not agree with more than seven of the ten classifications you should review the taxonomy before proceeding.

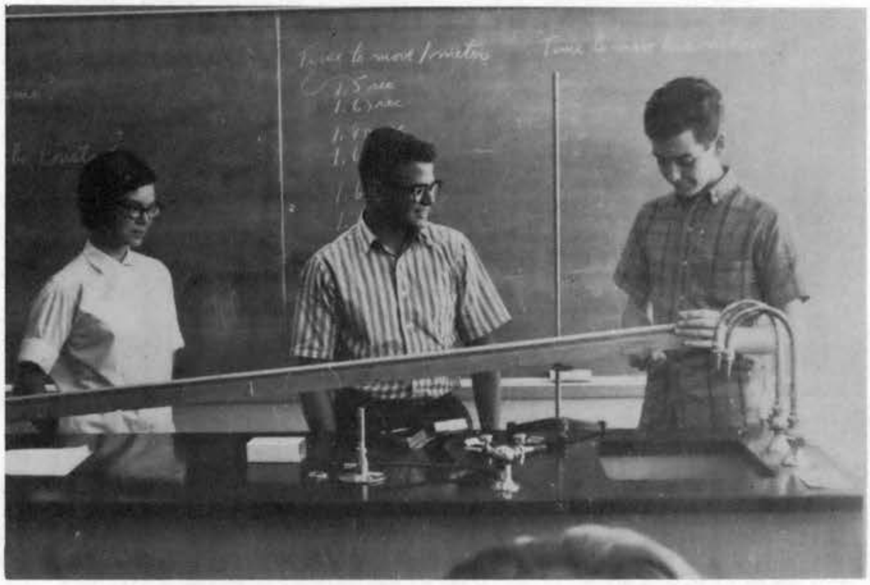
SCORING DIRECTIONS

You will be asked to classify thirty scenes of classroom behavior photographed from the hallway. You will see a series of 5 second film clips. At the designated times you will be asked to classify the

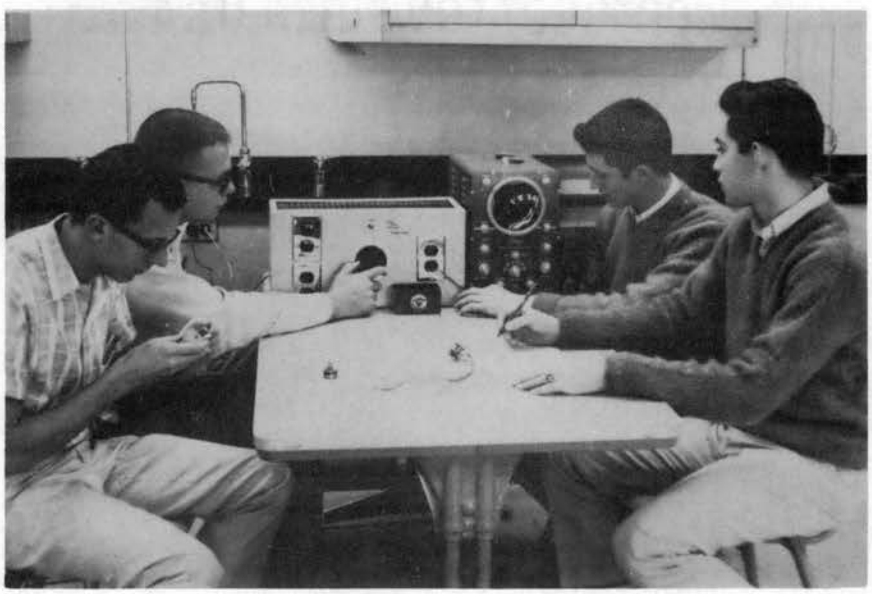
Picture No. 1



Picture No. 2



Picture No. 3



Picture No. 4

Picture No. 5

Picture No. 6

Picture No. 7

Picture No. 8

Picture No. 9

Picture No. 10

behavior at the moment. You should place a check mark in the space directly across from the category selected and in the column for that numbered observation. Example: Observation number 7 is classified 1.32.

	Observation No. 1	2	3	4	5	6	7	8	9
Code 1.11							.		
1.12							.		
1.13							.		
1.21							.		
1.22							.		
1.23							.		
1.31							.		
1.32							X		
1.33									
.									
.									

You will find a condensed form of the taxonomy on the inside front cover of your folder and the observers coding sheet on the inside back cover. You do not have to memorize the taxonomy. You simply remove the contents of your folder and follow the classifications across to the coding sheet. Your supervisor will clarify this point before observations begin. It is essential that we have your signature on the observation sheet.

ANSWERS: (A) 1.21 (B) 1.13 (C) 2.12 (D) 2.21 (E) 1.22 (F) 2.23

APPENDIX B

RECORDING FORMS FOR THE TAXONOMY OF CLASSROOM ACTIVITY

TAXONOMY OF CLASSROOM ACTIVITY

OBSERVERS RECORD SHEET NAME______________________

Listening & Viewing	Listening	To a mechanical device To a teacher To a student
	Viewing	An object or reproduction A teacher demonstration A student demonstration
	Listening-Viewing	A. V. Apparatus Listening to teacher & viewing Listening to student & viewing
Reading & Writing	Reading	Textbooks Non-textbooks
	Writing	Workbook or manual Student paper or notebook Testing (essay)
	Mixed Reading & Writing	Text, notebook, &/or workbook Objective testing Other reading & writing
Manipulative Tasks	Groups of 3	Using similar materials Using different materials Class out of classroom
	Groups of 2	Using similar materials Using different materials Unclassified
	Individual Work	Using similar materials Using different materials
Students Performing Two or More Tasks	Work Groups of Two or More	Activity involving 1, 2, & 3 Activity involving 1 & 2 Activity involving 1 & 3 Activity involving 2 & 3
	Individual Work	Activity involving 1, 2, & 3 Activity involving 1 & 2 Activity involving 1 & 3 Activity involving 2 & 3
	Individual and Group Work	Activity involving 1, 2, & 3 Activity involving 1 & 2 Activity involving 1 & 3 Activity involving 2 & 3

TAXONOMY OF CLASSROOM ACTIVITY OBSERVERS RECORD SHEET

DATE OF OBSERVATION_______________ NAME______________________________

SCENE NO.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
//////////	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///
//////////	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///
//////////	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///
//////////	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///
//////////	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///
//////////	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///
//////////	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///
//////////	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///
//////////	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///
//////////	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///
//////////	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///	///

APPENDIX C

RAW SCORES DATA

RAW SCORES DATA

SUM OF SIX OBSERVATION SCORES PER TEACHER

LEVEL I

Teachers	Observers Number 1	2	3	4	5	6	Total
C-1	94	93	83	101	76	97	544
C-2	77	121	99	107	75	96	575
C-3	41	83	52	100	76	86	438
C-4	113	85	91	110	75	128	602
C-5	79	102	60	78	104	74	497
S-1	53	60	54	79	72	74	392
S-2	142	112	161	167	139	139	860
S-3	126	96	92	67	126	104	611
S-4	41	62	18	41	34	40	236
S-5	136	128	107	151	99	126	747
Observer Total	902	942	817	1001	876	964	

$Y^2 = 567{,}089$

$\frac{Y^2}{n} = 504{,}533.4000$

$(\text{Totals})^2 = 3{,}308{,}108$

$(\text{Observer Totals})^2 = 5{,}067{,}130$

$Y = 5{,}502$

RAW SCORES DATA

SUM OF SIX OBSERVATION SCORES PER TEACHER

LEVEL II

Teachers	Observers Number 1	2	3	4	5	6	Total
C-1	37	34	31	36	30	35	203
C-2	29	44	37	38	27	36	211
C-3	16	30	20	36	29	31	162
C-4	41	32	31	38	29	46	217
C-5	30	37	24	30	39	29	189
S-1	20	23	21	28	27	28	147
S-2	53	42	57	60	52	51	315
S-3	47	34	35	26	45	38	225
S-4	16	23	8	16	13	16	92
S-5	50	46	39	54	38	45	272
Observer Total	339	345	303	362	329	355	

$Y^2 = 76,531$

$\frac{Y^2}{n} = 68,884.8166$

$(\text{Totals})^2 = 74,781.8333$

$(\text{Observer Totals})^2 = 691,065.0$

$Y = 2,033$

RAW SCORES DATA

SUM OF SIX OBSERVATION SCORES PER TEACHER

LEVEL III

Teachers	Observers Number 1	2	3	4	5	6	Total
C-1	15	13	12	14	13	14	81
C-2	7	17	14	14	11	13	76
C-3	9	12	9	15	11	12	68
C-4	15	12	12	12	7	17	75
C-5	12	14	11	12	14	12	75
S-1	10	10	10	12	7	10	59
S-2	19	16	20	22	19	18	114
S-3	18	14	13	11	14	14	84
S-4	9	9	6	9	6	6	45
S-5	18	16	14	20	16	16	100
Observer Total	132	133	121	141	118	132	

$Y^2 = 10,875$

$\frac{Y^2}{n} = 10,062.1500$

$(\text{Totals})^2 = 63,769$

$(\text{Observer Totals})^2 = 100,983$

$Y = 777$

RAW SCORES DATA

SUM OF SIX OBSERVATION SCORES PER TEACHER

LEVEL I

Teachers	Observers Number 7	8	9	10	11	12	Total
C-1	68	64	79	56	80	81	428
C-2	110	80	59	141	63	92	545
C-3	66	49	46	44	72	64	341
C-4	56	44	100	55	77	49	381
C-5	85	74	69	63	86	74	451
S-1	52	57	87	51	60	57	364
S-2	135	139	137	158	147	89	805
S-3	82	85	86	92	79	85	509
S-4	47	18	18	41	18	71	213
S-5	115	139	123	123	134	128	762
Observer Total	816	749	804	824	816	790	

$Y^2 = 450,275$

$\frac{Y^2}{n} = 383,840.0166$

$(\text{Totals})^2 = 435,111.1666$

$(\text{Observer Totals})^2 = 3,842,205$

$Y = 762$

RAW SCORES DATA

SUM OF SIX OBSERVATION SCORES PER TEACHER

LEVEL II

Teachers	Observers Number 7	8	9	10	11	12	Total
C-1	26	26	31	24	30	32	169
C-2	41	30	23	50	24	34	202
C-3	25	19	18	17	27	24	130
C-4	21	17	37	21	31	19	146
C-5	32	29	27	25	33	29	175
S-1	20	22	32	20	23	22	139
S-2	51	52	51	57	54	33	298
S-3	31	32	32	34	30	31	190
S-4	18	8	8	16	8	27	85
S-5	43	51	46	46	49	46	281
Observer Total	308	286	305	310	309	297	

$Y^2 = 63,277$

$\frac{Y^2}{n} = 54,903.7500$

$(\text{Totals})^2 = 368,617$

$(\text{Observer Totals})^2 = 549,475$

$Y = 1,815$

RAW SCORES DATA

SUM OF SIX OBSERVATION SCORES PER TEACHER

LEVEL III

Teachers	Observers Number 7	8	9	10	11	12	Total
C-1	10	12	13	10	13	13	71
C-2	15	12	10	19	9	13	78
C-3	10	7	7	7	10	9	50
C-4	9	8	14	9	12	9	61
C-5	12	12	11	11	13	12	71
S-1	10	10	13	10	10	10	63
S-2	18	18	18	20	19	12	105
S-3	12	12	12	13	11	12	72
S-4	18	6	6	9	6	11	56
S-5	15	18	17	16	18	17	101
Observer Total	129	115	121	124	121	118	

$Y^2 = 9,602$

$\frac{Y^2}{n} = 8,833.0666$

$(\text{Totals})^2 = 55,902$

$(\text{Observer Totals})^2 = 88,448$

$Y = 728$

RAW SCORES DATA

SUM OF SIX OBSERVATION SCORES PER TEACHER

LEVEL I

Teachers	Observers Number						Total
	13	14	15	16	17	18	
C-1	102	72	163	65	83	67	552
C-2	58	134	84	91	143	112	622
C-3	72	44	46	89	44	49	344
C-4	98	32	89	85	75	77	456
C-5	57	75	74	75	73	73	427
S-1	69	146	98	63	143	57	576
S-2	82	157	127	138	136	119	759
S-3	79	93	81	94	85	95	527
S-4	24	12	30	24	75	20	185
S-5	124	84	138	111	134	136	727
Observer Total	765	849	930	835	991	805	

$Y^2 = 524,423$

$\frac{Y^2}{n} = 446,343.7500$

$(\text{Totals})^2 = 2,948,529$

$(\text{Observer Totals})^2 = 4,498,257$

$Y = 5,175$

RAW SCORES DATA

SUM OF SIX OBSERVATION SCORES PER TEACHER

LEVEL II

Teachers	Observers Number 13	14	15	16	17	18	Total
C-1	40	37	31	26	31	26	191
C-2	22	48	32	34	51	40	227
C-3	27	17	18	33	17	19	131
C-4	36	13	33	31	28	29	170
C-5	23	29	29	29	29	29	168
S-1	26	53	36	24	52	22	213
S-2	32	57	47	51	51	44	282
S-3	30	35	31	35	32	35	198
S-4	10	6	12	10	28	9	75
S-5	45	32	51	41	49	50	268
Observer Total	291	327	320	314	368	303	

$Y^2 = 70,569$

$\frac{Y^2}{n} = 61,632.1500$

$(\text{Totals})^2 = 403,841$

$(\text{Observer Totals})^2 = 619,839$

$Y = 1,923$

RAW SCORES DATA

SUM OF SIX OBSERVATION SCORES PER TEACHER

LEVEL III

Teachers	Observers Number 13	14	15	16	17	18	Total
C-1	16	15	13	12	12	10	78
C-2	9	18	12	12	18	15	84
C-3	10	7	7	12	7	7	50
C-4	14	8	12	13	11	12	70
C-5	10	12	12	12	12	12	70
S-1	10	20	13	10	19	10	82
S-2	11	20	17	18	18	16	100
S-3	11	13	12	13	13	13	75
S-4	6	6	6	6	11	6	41
S-5	17	12	18	15	18	18	98
Observer Total	114	131	122	123	139	119	

$Y^2 = 10,194$

$\frac{Y^2}{n} = 9,325.0667$

$(\text{Totals})^2 = 59,074$

$(\text{Observer Totals})^2 = 93,652$

$Y = 748$

RAW SCORES DATA

SUM OF SIX OBSERVATION SCORES PER TEACHER

LEVEL I

Teachers	Observers Number						Total
	19	20	21	22	23	24	
C-1	54	94	95	79	80	37	439
C-2	130	52	51	117	150	62	562
C-3	100	60	67	40	31	69	367
C-4	81	103	62	131	89	65	531
C-5	106	72	70	73	106	50	477
S-1	62	55	141	72	79	135	544
S-2	145	121	100	85	137	98	686
S-3	155	89	124	103	115	65	651
S-4	36	27	98	61	109	71	402
S-5	176	88	116	76	168	97	721
Observer Total	1,045	761	924	837	1,064	749	

$Y^2 = 544,044$

$\frac{Y^2}{n} = 482,406.6667$

$(\text{Totals})^2 = 3,024,522$

$(\text{Observer Totals})^2 = 4,918,588$

$Y = 5,380$

RAW SCORES DATA

SUM OF SIX OBSERVATION SCORES PER TEACHER

LEVEL II

Teachers	Observers Number 19	20	21	22	23	24	Total
C-1	21	35	35	29	31	15	166
C-2	45	24	19	43	52	23	206
C-3	37	24	25	16	13	26	141
C-4	31	39	24	48	32	27	201
C-5	38	25	26	27	38	20	174
S-1	23	21	51	26	30	48	199
S-2	52	45	39	33	49	37	255
S-3	54	34	45	38	42	25	238
S-4	14	11	36	22	39	27	149
S-5	63	34	43	29	60	36	265
Observer Total	378	292	343	311	386	284	

$Y^2 = 74,614$

$\frac{Y^2}{n} = 66,267.2667$

$(\text{Totals})^2 = 414,246$

$(\text{Observer Totals})^2 = 672,170$

$Y = 1,994$

RAW SCORES DATA

SUM OF SIX OBSERVATION SCORES PER TEACHER

LEVEL III

Teachers	Observers Number 19	20	21	22	23	24	Total
C-1	9	14	13	11	12	8	67
C-2	16	11	7	16	19	9	78
C-3	14	10	10	8	6	10	58
C-4	13	16	10	18	12	11	80
C-5	15	12	11	11	14	9	72
S-1	10	10	20	11	13	18	82
S-2	20	18	15	13	19	15	100
S-3	20	13	17	14	16	10	90
S-4	10	6	14	9	15	10	64
S-5	23	14	16	13	21	16	103
Observer Total	150	124	133	124	147	116	

$Y^2 = 11,410$

$\frac{Y^2}{n} = 10,507.2667$

$(\text{Totals})^2 = 65,050$

$(\text{Observer Totals})^2 = 106,006$

$Y = 794$

RAW SCORES DATA

SUM OF SIX OBSERVATION SCORES PER TEACHER

LEVEL I

Teachers	Observers Number 25	26	27	28	29	30	Total
C-1	84	138	113	69	68	133	605
C-2	113	75	65	71	101	136	561
C-3	63	54	73	72	43	98	403
C-4	81	69	107	103	91	60	511
C-5	73	70	103	107	82	103	538
S-1	64	108	118	51	80	121	542
S-2	156	149	101	139	96	148	789
S-3	62	147	94	68	76	124	571
S-4	68	54	43	29	77	106	377
S-5	105	126	115	134	88	174	742
Observer Total	869	990	932	843	802	1,203	

$Y^2 = 590,099$

$\frac{Y^2}{n} = 529,972.0167$

$(\text{Totals})^2 = 3,328,739$

$(\text{Observer Totals})^2 = 5,404,947$

$Y = 5,639$

RAW SCORES DATA

SUM OF SIX OBSERVATION SCORES PER TEACHER

LEVEL II

Teachers	Observers Number 25	26	27	28	29	30	Total
C-1	32	51	42	27	27	48	227
C-2	41	28	25	26	36	48	204
C-3	24	21	27	27	17	37	153
C-4	30	26	39	38	33	22	188
C-5	30	27	39	39	31	38	204
S-1	23	40	43	20	30	44	200
S-2	59	53	38	51	36	54	291
S-3	25	52	35	26	29	44	211
S-4	25	21	17	12	29	39	143
S-5	38	46	43	49	33	61	270
Observer Total	327	365	348	315	301	435	

$Y^2 = 80,089$

$\frac{Y^2}{n} = 72,871.3500$

$(\text{Totals})^2 = 456,065$

$(\text{Observer Totals})^2 = 740,309$

$Y = 2,091$

RAW SCORES DATA

SUM OF SIX OBSERVATION SCORES PER TEACHER

LEVEL III

	Observers Number						
Teachers	25	26	27	28	29	30	Total
C-1	13	20	17	11	11	18	90
C-2	15	11	10	10	13	17	76
C-3	10	10	9	11	8	14	62
C-4	12	10	16	15	13	10	76
C-5	12	11	15	15	12	15	80
S-1	10	16	16	10	13	16	81
S-2	21	20	15	18	14	19	107
S-3	11	19	13	11	12	16	82
S-4	9	9	9	6	13	14	60
S-5	14	18	16	18	13	22	101
Observer Total	127	144	136	125	122	161	

$Y^2 = 11,821$

$\frac{Y^2}{n} = 11,070.4167$

$(\text{Totals})^2 = 68,431$

$(\text{Observer Totals})^2 = 111,791$

$Y = 815$

RAW SCORES DATA

SUM OF SIX OBSERVATION SCORES PER TEACHER

LEVEL I

Teachers	Observers Number						Total
	31	32	33	34	35	36	
C-1	99	89	85	69	80	93	515
C-2	121	96	159	140	107	82	705
C-3	26	71	106	54	57	66	380
C-4	62	98	128	66	112	83	549
C-5	71	90	79	77	72	79	468
S-1	137	91	51	72	77	74	502
S-2	155	106	147	143	155	138	844
S-3	83	88	143	94	92	90	590
S-4	12	68	105	14	64	36	299
S-5	133	110	148	115	130	106	742
Observer Total	899	907	1,151	844	946	847	

$Y^2 = 591,044$

$\frac{Y^2}{n} = 521,547.2666$

$(\text{Totals})^2 = 3,379,480$

$(\text{Observer Totals})^2 = 5,280,312$

$Y = 5,594$

RAW SCORES DATA

SUM OF SIX OBSERVATION SCORES PER TEACHER

LEVEL II

Teachers	Observers Number						Total
	31	32	33	34	35	36	
C-1	37	33	31	26	31	35	193
C-2	44	35	57	50	39	31	256
C-3	11	27	39	21	21	25	144
C-4	23	37	46	25	40	30	201
C-5	29	34	30	29	29	29	180
S-1	50	33	20	27	29	28	187
S-2	57	40	54	53	55	50	309
S-3	31	32	47	35	35	33	213
S-4	6	26	38	7	24	14	115
S-5	48	41	54	42	47	39	271
Observer Total	336	338	416	315	350	314	

$Y^2 = 79,657$

$\frac{Y^2}{n} = 71,346.0166$

$(\text{Totals})^2 = 458,807$

$(\text{Observer Totals})^2 = 720,512$

$Y = 2,069$

RAW SCORES DATA

SUM OF SIX OBSERVATION SCORES PER TEACHER

LEVEL III

Teachers	Observers Number						Total
	31	32	33	34	35	36	
C-1	15	13	13	11	13	13	78
C-2	16	13	20	18	15	11	93
C-3	7	10	14	9	7	9	56
C-4	10	14	13	11	15	11	74
C-5	12	13	12	12	12	12	73
S-1	19	12	10	12	11	10	74
S-2	20	15	19	19	20	18	111
S-3	13	12	17	13	14	12	81
S-4	6	10	15	6	9	6	52
S-5	18	15	19	16	18	14	100
Observer Total	136	127	152	127	134	116	

$Y^2 = 11,246$

$\frac{Y^2}{n} = 10,454.0000$

$(\text{Totals})^2 = 65,736$

$(\text{Observer Totals})^2 = 105,270$

$Y = 792$

VITA

Neal Jay Holmes

Candidate for the Degree of

Doctor of Education

Thesis: THE DEVELOPMENT AND EVALUATION OF AN OBSERVATIONAL SYSTEM DESIGNED TO MEASURE SCIENCE CLASSROOM ACTIVITY FROM THE HALL

Major Field: Higher Education

Biographical:

Personal Data: Born at Mercer, Missouri, August 2, 1931, the son of Allie O. and Marie Holmes. Married Mary Lou Liggett, December 20, 1953; four sons, Tommie L., Timothy A., Theodore J., and Todd Andrew.

Education: Attended grade school in Mercer County, Missouri, and Wayne County, Iowa; graduated from Allerton, Iowa High School in 1949; received the degree of Bachelor of Science in Education from Northeast Missouri State Teachers College in 1957. Graduate Study in Science Education summers of 1958-59, Northeast Missouri State Teachers College. Graduate Study in chemistry and mathematics summers of 1960-61, Missouri School of Mines. Received the degree of Master of Arts from Washington University, St. Louis, Missouri, in 1962.

Professional Experience: Served with U.S. Air Force from 1952-1956 with three years' service in Europe. Taught science and mathematics in high school in Washington, Missouri, from 1957-1961. Served as Science Consultant K-12 for the Parkway School District, St. Louis County, Missouri, from 1962 to present time.

Professional Organizations: Phi Delta Kappa, N.E.A., N.S.T.A., N.S.S.A., M.S.T.A., S.L.S.F.A., Missouri Science Teachers Association.