DOCUMENT RESUME

ED 128 339 SP 010 411

AUTHOR Berliner, David C.
TITLE The Beginning Teacher Evaluation Study: Overview and Selected Findings, 1974-1975.
INSTITUTION Far West Lab. for Educational Research and Development, Berkeley, Calif.
PUB DATE Nov 75
NOTE 29p.; Paper presented at the National Invitational Conference on Research on Teacher Effects: An Examination by Decision-Makers and Researchers (Austin, Texas, November 3-5, 1975)

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS *Educational Research; *Effective Teaching; Elementary Education; *Elementary School Teachers; *Ethnology; Mathematics; *Performance Criteria; Rating Scales; Reading; Teacher Education; *Teacher Evaluation; Teacher Improvement
IDENTIFIERS California

ABSTRACT

The study presented here examines whether an ethnographic approach to the study of teaching yields new insight into the teaching-learning process. Two-hundred teachers, who differed in measured effectiveness, were recruited from thirteen school districts in the state of California. Each teacher taught two experimental teaching units (ETU's) of two-weeks duration. Each ETU included an introduction to the teacher, giving a rationale for the unit; performance objectives; pre- and post-examinations for students; and a variety of instructional materials and activities. After test data were collected, posttest scores were regressed on pretest scores for each grade level. Based on class pretest means, three strata were created: low, middle, and high-achieving classrooms within each subject area and grade level. Twelve observers were selected and trained. This included learning to read educational ethnographies, practicing in classrooms, and observing films of classrooms. The ethnographers were trained to provide both reading and mathematics protocols each day; give informal protocols based on observations during recess, talks with principals, and conversations with peers; and asked to give a summary protocol emphasizing important anthropological concepts useful for studying education. Six raters were brought together for two weeks to read a pair of protocols a day, describing a more effective and less effective classroom. They were asked to describe as many ways as possible that the two classrooms differed using any desired terminology. They generated 211 dimensions. This list was revised to 61 variables and used to do a more extensive study involving 20 raters using specially constructed rating forms. (DMT)

THE BEGINNING TEACHER EVALUATION STUDY: OVERVIEW AND SELECTED FINDINGS, 1974-1976

DAVID C. BERLINER

Far West Laboratory for Educational Research and Development

November, 1975

Paper Presented at the University of Texas, Austin, Texas, November 2-4, 1975, for the Conference on Research on Teacher Effects: An Examination by Decision-Makers and Researchers.

INTRODUCTION

The Far West Laboratory for Educational Research and Development has been conducting research on teacher effectiveness. Under contract to the California Commission for Teacher Preparation and Licensing, with funding provided by the National Institute of Education, the Laboratory is studying teachers in second- and fifth-grade classes in order to identify teacher behavior and classroom qualities that are related to reading and mathematics achievement.

The California Commission is the agency charged with certifying the appropriateness of teacher training programs throughout the state. To carry out its duties, the Commission needs information about what teacher behaviors are related to student outcomes. This information will then be used jointly by the Commission and the State institutions that it certifies in order to better insure that beginning teachers receive training in areas that have been empirically demonstrated to affect student learning.

To obtain the information they need, the Commission has undertaken a multi-year research effort entitled the Beginning Teacher Evaluation Study (BTES). During 1974-1975, as part of this study, the Laboratory did work on five major tasks. Three of these were substantive tasks, inquiring: 1) whether ethnographic approaches to the study of teaching could yield new insights into the teaching-learning process; 2) whether the planning and decision-making engaged in by teachers is different for teachers who vary in their ability to induce learning; and 3) whether the perceptions of teachers who vary in their ability to induce learning differ when viewing teaching-learning situations, and, of equal interest, whether students'

perceptions of the teaching-learning situation would yield information about what aspects of classroom interaction were salient and important to the students.

Two methodological issues were also addressed during 1974-1975. One of these issues was concerned with the explication of a conceptual model, with methodological procedures, in the area of time allocation in natural classroom situations. It has become increasingly evident to us that instructional time is related to educational outcomes. The BTES staff is continuing to investigate this area. The second methodological issue was concerned with an inquiry into generalizability theory applied to the problem of how many observers and/or how many occasions are necessary before stable estimates of a teacher's classroom behavior are recorded.

The goal of these five tasks, as well as some others, was to provide information that could be combined with existing data obtained by Educational Testing Service during 1973-1974, in order to design a large field study examining teacher effectiveness. I shall describe in detail only one of our efforts, the ethnographic study. Information about other aspects of the study can be obtained from the Far West Laboratory by requesting the BTES Technical Reports.

Identification of a Sample of More Effective and Less Effective Teachers

A first step in our approach to generate variables worthy of further consideration in the study of teacher effectiveness was to identify a sample of teachers in the State of California who differed in their measured effectiveness. Two hundred teachers were recruited from thirteen school districts in the State of California. Half were second grade

teachers and half were fifth grade teachers. Each teacher taught two Experimental Teaching Units (ETU's) of two-week's duration each. The second grade ETU in reading centered on using word structure as a tool for comprehension; reading for comprehension and content; following directions; and sequencing. The second grade ETU in mathematics focused on the concept of measurement, conservation, and the associated arithmetic operations.

T[illegible] reading focused on reading in larger units (phras[illegible] The fifth grade ETU in mathematics cent[illegible] d associated arithmetic operations.

[illegible] the teacher giving a rationale for the [illegible]onna[illegible] t were tied to the tests for the unit; pretests for the stude[illegible]ctional materials and activities for the teacher to choose from, and a posttest for the students. The ETU's conformed to accepted state curriculum objectives for grades two and five, but were in areas not ordinarily stressed by teachers in those grades. In the second grade, instruction took about twenty minutes a day for two weeks, and in the fifth grade instruction took about forty-five minutes a day for two weeks. The ETU's in the two subject matters were administered sequentially.

After the test data were collected, classroom posttest scores were regressed on pretest scores for each ETU, separately for each grade level. Based on class pretest means, three strata were created so that the regressions were run separately for low, middle, and high achieving classrooms within subject matter area, and grade level. From the initial sample of one hundred teachers at each grade level, ten classrooms with high residual gain scores in reading, mathematics, or both areas, were chosen.

Ten classrooms with low residual gain scores in reading, mathematics, or both areas were also chosen. These classrooms came from all three pretest strata. The ten teachers with classrooms that showed higher than predicted posttest means were designated as "more effective" teachers. The ten teachers associated with classrooms that showed lower than predicted posttest means were designated as "less effective" teachers. It must be noted that these designations are relative and that the teachers in the study were a volunteer sample. Our procedures can only allow us to talk of more and less effective teachers, as measured in the manner described, and in no way implies that these teachers are "good" or "bad" in their regular classroom performance.

The pretest and posttest correlations for the ETU's, in both subject areas, and both grades, were rather high (around .90). Correlations between residuals across curriculum areas were low, around .30, in both grades. In correlating the ETU gains in reading and mathematics with the gains from year-long testing using the California Achievement Test and a special battery of tests created by the Educational Testing Service, it was found that the ETU gain scores over two weeks were positively related to these other more standard measures of gain over the entire year.

The ETU's and associated tests were designed to identify a sample of teachers who varied in measured effectiveness when teaching a common curriculum, to common objectives, for controlled amounts of time. On the basis of residual gain scores over the two weeks of teaching each ETU, ten more and ten less effective teachers at each grade level were identified and constituted the known sample of teachers that was used for intensive analysis of classroom practice. The ethnographic study I am reporting on today was conducted with this special sample of teachers.

Ethnography in the Classroom*

Our goal was to obtain protocols of classrooms written by sensitive observers who were unaware of the measured effectiveness of the teachers they observed. The BTES staff believed that "single-act" psychology and hypothesis testing psychology had yielded little of value for studying the complex world of the classroom. Thus we turned elsewhere for a way of viewing classroom phenomena. Recently, our own feelings of uneasiness with traditional psychological approaches have been echoed by others. Lutz and Ramsey (1974) have been concerned that the teaching acts and learning outcomes that have been studied to date are those that, for the most part, can be subjected to measurement by paper and pencil tests and/or by the development of behaviorally defined coding systems. Descriptions of the activity in a classroom, therefore, have been limited by the "screens" through which events have been recorded--those "screens" being soundly based from a psychometric quantitative point of view, but lacking in terms of qualitative information surrounding the reality of what actually occurred. They say:

> Variables are operationalized because there is some available printed test with some kind of statistical reliability and validity measure, and after data are collected, it can be submitted to a computer for an analysis usually much too esoteric and powerful for the nature of the hypothesis. In such a case, the hypothesis is not grounded, the variables may not be recurring or important; the operational measures may have little relationship to operational reality, and the number in the sample, make test of it much more powerful than the hypothesis is compelling (Lutz and Ramsey, 1974, p.5)

---

*The research to be described was directed by Dr. William Tikunoff. A complete report of these activities is given in Tikunoff, W., Berliner, D.C. and Rist, R.C. An ethnographic study of the forty classrooms of the Beginning Teacher Evaluation Study known sample. Technical Report No. 75-10-5 San Francisco, Calif.: Far West Laboratory for Educational Research and Development, October, 1975.

The result of the shortcomings from research on teaching, and the uneasiness shared by many with the intellectual style of psychological research has established an introspectic stance by some educational researchers toward their accomplishments to date (Campbell, 1974; Cronbach, 1975; Glass, 1972). Such reflection has led to an intensive questioning of the research questions which are asked, and therefore, the research methodology being employed to answer them. It is out of that questioning that the impetus has grown to look beyond the methodologies of experimental psychology to other disciplines of the social sciences for the purpose of studying teaching. Particularly important is anthropology or social anthropology and their observational techniques. The use of the direct observer, fully imbedded in the on-going process of the classroom, seems to be an emerging tool for use in some current evaluations of NIE-funded projects (Campbell, 1974), and the use of such anthropological field methodology over a longer period of time should result in accumulation of more qualitative data of, potentially, great utility (Lutz and Ramsey, 1974). Our goal must be to gather more qualitative information along with the quantitative information we usually collect. Campbell (1974) characterizes the contrast between these two approaches:

> For quantitative read also scientific, scientistic and naturwissenschaftlich. For qualitative read also humanistic, humanitistic, geisteswissenschaftlich, experiential, phenomenological, clinical, case study, field work, participant observations, process evaluation, and common-sense knowing.

The gathering of such qualitative evidence, suggests Cronbach, involves intensive local observation that goes beyond disciplines to an open-eyed, open-minded appreciation of the surprises nature deposits in the investigative act (Cronbach, 1975). It necessitates the "direct observation of

human activity and interaction in an ongoing, naturalistic fashion" (Rist, 1973). It allows the researcher to:

> File descriptive information ... instead of reporting only those selected differences and correlations that are nominally "greater than chance." (Cronbach, 1975).

It was because of these concerns and beliefs that an ethnographic study was designed and conducted.

Recruitment of Ethnographers

The first task was the recruitment of sensitive observers to send to the specially selected classrooms to obtain the qualitative information that was desired. The 12 observers finally chosen are described in Table 1. Most were doctoral candidates, in anthropology or sociology, and most had served as nonparticipant or participant observers previously.

- - - - - - - - - - - -
Insert Table 1 Here
- - - - - - - - - - - -

Training of these ethnographers consisted of 1) reading educational ethnographies, including those of Jules Henry and others; 2) practice in classrooms; and 3) observing films of classrooms. Protocols produced during training were read and critiqued by the Laboratory staff. Three weeks of effort was directed into getting the ethnographers ready to focus on reading and mathematics lessons in natural classrooms. A sample protocol is given as Figure 1 and should be read carefully to give you a feel for the kind of data we collected and worked with. This particular sample is a training protocol which was read and critiqued by the Laboratory staff during the time the ethnographers were being trained.

- - - - - - - - - - - -
Insert Figure 1 Here
- - - - - - - - - - - -

Although we were interested in obtaining qualitative information, we did not ignore the chance to "calibrate" our data collectors. Table 2 presents data on that issue. An expert ethnographer was used as criterion during some training exercises. From a film clip of a classroom discussion he picked nine events that were salient to him. The ethnographers saw the same film clip and wrote protocols on what they observed. The information in Table 2 on percent agreement with the expert is to be interpreted as a form of a validity check. The information on percent agreement among the raters is to be interpreted as a form of a reliability check. All disagreements were used for discussions about recording observations. When reliability and validity were judged high enough, these observers were sent out into the field.

- - - - - - - - - - -
Insert Table 2 Here
- - - - - - - - - - -

Data Base

The ethnographers were trained to provide:

1) A reading protocol, each day, if reading was taught;
2) A mathematics protocol, each day, if mathematics was taught;
3) Three to five informal protocols based on observations during recess, talks with principals, conversations in the teacher's lounge, etc.;
4) A summary protocol emphasizing important anthropological concepts useful for studying education. These concepts include competitiveness, work ethic, patriotism, play ethic, etc.

Thus the data set, with one ethnographer observing for one week in each classroom includes five reading protocols, five mathematics protocols,

at least three informal protocols, and one summary protocol done after observation was completed. Four weeks were required to collect data from all the teachers in the study. All data were collected blind. The ethnographers had, typically, two more effective and two less effective teachers to work with and worked only at one grade level. Classroom notes were read into a cassette recorder each day and sent to the Laboratory for immediate transcription. Teams of typists helped to turn out thousands of pages describing classrooms of these teachers who were known to vary in measured effectiveness.

At this point you now have knowledge of how this particular sample of teachers were chosen, what the goal of the study was, what the training of the ethnographers was like, and a description of the data sets obtained in each of forty classrooms.

Generating Dimensions

Six raters were brought together for two weeks to read a pair of protocols a day. One protocol described a more effective classroom and one protocol described a less effective classroom. These raters included one expert in classroom observation instrumentation, a classroom teacher, a curriculum coordinator, one graduate student in educational psychology, and two ethnographers who were thought to write very sensitive descriptive protocols. The raters were asked to describe as many ways as possible that the two classrooms differed. They were free to use any terminology they wanted. They were aware that they had a more effective and a less effective classroom paired together, but they did not know which classroom was which. The hope was to keep this task relatively hypothesis free and at a common sense level. To help them in their task the raters used cards like those presented in Figure 2. This task was, essentially, a concept-definition

task. At the end of each day, the raters came together and shared their concepts. Each rater helped other raters define the concepts and each provided exemplars and non-exemplars of the concepts from their own protocols. The list generated by these raters contained 211 concepts. These concepts are presented in Table 3.

- - - - - - - - - - - -
Insert Figure 2 Here
- - - - - - - - - - - -
- - - - - - - - - - - -
Insert Table 3 Here
- - - - - - - - - - - -

Remember, that each of these sometimes exotic variables, dimensions or concepts was linguistically defined in a rather precise manner. Thus concepts like "psychotic autism," or a view that "children are evil," were concepts that we had no preconceived desire to work with. But they were chosen as concepts that differentiated between more and less effective classrooms by at least one of the raters and the concepts were agreed to, refined, and defined, by the other raters. We purposefully did not place any limits on the type of concepts that could be generated. To this list eight additional concepts were added. These included five concepts from Kounin's (1970) work (withitness, smoothness, transitions, etc.), and three variables that were experimentally manipulated in a study conducted at the Stanford Research and Development Center. It was thought that an independent correlational check of those variables could be made in this study.

The list of 211 dimensions was much too big to work with and contained a good deal of overlapping concepts. The dimensions were combined into 61 variables which were thought to capture most of what the "dimension pickers" had chosen. Variables were also chosen on the basis of whether or not they appeared frequently in the protocols. The final list of 61

dimensions, and brief definitions, is given as Table 4. A "T" or "S" denotes the variable as related to the teacher or the student as the focus of observation. These 61 variables were next used to do more extensive analysis of the protocols.

- - - - - - - - - - -
Insert Table 4 Here
- - - - - - - - - - -

Twenty raters from all walks of life were brought together to rate pairs of protocols for the presence or absence, or the occurrence of more or less of the variable. They received training in the use of a specially constructed rating form and in understanding the definitions for variables. The rating forms used in conjunction with the manual defining each of the 61 dimensions is given as Figure 3. A sample of the rating manual is given as Table 5. Each variable is clearly defined and examples of each variable are taken directly from the protocols that describe natural classroom behavior.

- - - - - - - - - - -
Insert Table 5 Here
- - - - - - - - - - -
- - - - - - - - - - - -
Insert Figure 3 Here
- - - - - - - - - - - -

Ten raters worked on second grade reading protocols and then were switched to work on fifth grade mathematics protocols. Ten other raters worked with fifth grade reading protocols and then were switched to second grade mathematics protocols. Each rater received a protocol of a more and a less effective teacher, according to the sampling plan presented as Figure 4. Each of the ten more effective teachers were treated as an interchangeable set and each of the ten less effective teachers were viewed as an interchangeable set. Pairs of classes were randomly picked for raters

to examine. Out of the total 100 pairs possible, 32 pairs were rated by one rater, and four additional pairs were rated independently by two raters. When this process was repeated for each category (reading and mathematics, second and fifth grade) we had sixteen reliability checks nested within the actual ratings. Within each grade and subject matter area 36 pairs of classrooms were compared. Each more effective classroom was compared three or four times with a less effective classroom.

- - - - - - - - - - - -
Insert Figure 4 Here
- - - - - - - - - - - -

Findings

Table 6 provides the summary data from this study. The simple binomial test was used to examine the ratings. Thirty-six opportunities for rating occurred, thus a split of eighteen and eighteen would have meant that the dimension, say "abruptness", was found eighteen times to be rated as occurring more often in the less effective classrooms and eighteen times to have occurred more often in the more effective classrooms A split by the raters of twenty-two and fourteen has a probability of occurrence of .09 and a split of twenty-three and thirteen has a probability of occurance of .05.

Insert Table 6 Here

This Table reveals that there are twenty-one variables that were generic. That is, these variables discriminated between more and less effective teachers in second grade reading, second grade mathematics, fifth grade reading, and fifth grade mathematics. Variables such as "teacher monitors learning (No. 37)," and " students are engaged (No. 19)," were consistently associated with the more effective teachers. A variable such as "teacher belittles students (No. 7)" was consistently found in the less effective teachers classrooms, regardless of the subject matter taught or the grade level examined.

Other variables were significantly associated with the more or less effective teachers within a single grade of a single subject matter area. And some variables were associated with effectiveness only in a particular grade x subject matter context. All 61 variables were significantly associated with the measured effectiveness of the teachers at least once when the various combinations of curriculum area and grade level were examined.

Conclusion

As our work continues each of these variables will be given closer examination in partial replications. In that way increased assurance about the validity of these variables for differentiating more and less effective teachers will be obtained. For now, this work meets the project goal which was to generate variables of promise in the study of teacher effectiveness. Both the methodology used and the results of this study are, in the opinion of the BTES staff, worth further investigation.

REFERENCES

Lutz F.W. and Ramsey, M.A. The use of anthropological field methods in education. Educational Researcher, 1974, 11, 5-9.

Campbell, D.T. Qualitative knowing in action research. Paper presented at the meeting of the American Psychological Association, New Orleans, Louisianna, September, 1974.

Cronbach, L.J. Beyond the two disciplines of scientific psychology. American Psychologist, 1975, 30, 116-127.

Glass, G.V. The wisdom of scientific inquiry on education. Journal of Research in Science Teaching, 1972, 9, 3-18.

Kounin, J.S. Discipline and group management in classrooms. New York: Holt, Rinehart and Winston, 1970.

TABLE 1

Background of Ethnographers.

Ethnographer	Most Recent School	Anthropology/ Sociology Degree Programs			Number of Nonparticipant/ Participant Observation Experiences
		Ph.D	M.A.	A.B.	
1.	U.C. Santa Barbara Ph.D, Sociology	X	X	X	3
2.	Stanford University M.A., Education			X	4
3.	Stanford University M.A., Communications	X*			0
4.	U.C. Berkeley Ph.D., Sociology of Education	X	X		2
5.	San Francisco State M.A., Sociology		X	X	2
6.	U.C. Davis Ph.D.* Sociology	X*		X	5
7.	U.C. Santa Barbara M.A., Sociology	X*	X	X	2
8.	U.C. Berkeley Ph.D, Sociology	X*		X	1
9.	U.C. San Francisco Ph.D*, Sociology	X*	X		1
10.	Stanford University M.A., Anthropology	X*	X		0
11.	Stanford University M.A., Anthropology	X*	X		1
12.	San Francisco State M.A., Sociology		X	X	3

*Ph.D. candidate

TABLE 2

Comparison of Ethnographers with the Trainer and with Each Other for Events Presented in the Film Clip During Training.

Events Observed in Film Clip	Ethnographers												% Agreement with Trainer by item
	1	2	3	4	5	6	7	8	9	10	11	12	
1.	X	X		X	X	X	X				X	X	.72
2.	X	X	X				X	X		X		X	.72
3.		X	X	X	X					X		X	.54
4.	X	X	X	X	X	X	X			X	X	X	.90
5.	X		X	X	X	X	X	X			X		.72
6.												X	.09
7.		X		X				X				X	.36
8.	X	X	X	X	X	X	X			X	X	X	.90
9.			X	X			X						.27
% Agreement Among Ethnographers	.55	.66	.66	.77	.55	.44	.66	.33		.44	.44	.77	

Mean agreement: .57

* Ethnographer 9 was absent due to illness.

TABLE 3

Teacher, Student, or Instructional Characteristics Generated by Raters

1. Transition/easing in
2. Concern for physical environment of classroom
3. Emphasis for private work
4. Complimenting as control
5. Promoting self-sufficiency
6. Explaining praise
7. Expressing optimism
8. Teacher feels lack of importance
9. Inconsistency in behavior
10. Teacher inattentiveness
11. Negative attitude toward curriculum materials
12. Goal clarification
13. Process orientation
14. Differential treatment by category/stereotyping
15. Depersonalization
16. Teacher expresses legitimate authority
17. Product orientation
18. Student self-defense
19. Emphasis on interpretation
20. Students express affection towards each other
21. Guilt/shame
22. Review/recall
23. Praising behavior
24. Use of affectional terms or affection
25. Smoothness
26. Peer teaching
27. The right to win
28. Internalization of classroom norms
29. Boredom disattention
30. Neglect of subject matter
31. Humorous spontaneity
32. Time rigid structure
33. Flexibility
34. Group alerting
35. Moralizing
36. Utilizing resources
37. Being liked
38. Monitoring learning
39. Overlapping
40. Withitness
41. Student rejecting/resisting
42. Pause/rhetorical questions
43. Competition
44. Beating
45. Gendering
46. Sense of shame
47. Illogical threat
48. Pointless teacher statements
49. Teacher non-verbal signal for controlling
50. Teacher autocracy
51. Social isolates
52. Relation of material to personal selves
53. Signals for order
54. Acceptance of feelings and attitudes
55. Protectiveness
56. Social closeness
57. Respect accommodation
58. Physical punishment (harrassing)
59. Flexibility
60. Expectancy (blaming the victim)
61. Prompting
62. Utilizing resources
63. Learning with variety
64. Spontaneity
65. Pacing
66. Rituals
67. Sarcasm
68. Hostility
69. Differential treatment of individual
70. Regimenting
71. Teacher authoritarianism
72. Teacher concern for student feelings
73. Reacting to student questions
74. Pursuing for thinking process
75. Discipline freak
76. Psychotic autisms
77. Guessing questions
78. Inter-student aggressiveness
79. Cooperativeness among students
80. Tacit permissiveness for student movement
81. Masking emotions
82. Socializing about obeying authority
83. Teacher encouraging speed
84. Inappropriate pacing
85. Inattentiveness to student needs
86. Distribution of responsibility
87. Lack of awareness of learning rates
88. Effective voice
89. Reasonableness of commands
90. Exclusion/isolation as control
91. Student defiance
92. Accepts own error/explanations not clear
93. Tactlessness
94. Loss of composure
95. Lack of mobility
96. Mixed messages
97. Conviviality
98. Modeling
99. Children are evil
100. Encouraging
101. Procedural flexibility
102. Distributes responsibility/shares responsibility
103. Teacher degradation of students
104. Teacher centrality

105. Structuring by teacher
106. Teacher changes activity abruptly
107. Teacher's time fixedness
108. Teacher lack of responsiveness
109. Teacher responsiveness
110. Teacher role clarity
111. Teacher constancy of message
112. Teacher modeling and student imitation
113. Teacher asks for different tasks according to sex
114. Adult standards for children
115. Pointless teacher statements
116. Teacher threats and ultimatums not carried out
117. Teacher threats and ultimatums carried out
118. Teacher lacks knowledge of subject area
119. Student's decoding teacher's intent
120. Disattention
121. Teacher's developmental awareness
122. Cross age tutoring
123. Teacher diagnosis and prescription
124. Teacher emphasized cooperation
125. Teacher accepted student's approaches
126. Teacher insecurity
127. Teacher's positive communication of race, class
128. Students used teacher's last name
129. Mobility of teacher
130. Teacher handles hostility
131. Teacher's emphasis on orderliness
132. Teacher's emphasis on cleanliness
133. Value conflict
134. Teacher's positive rewards
135. Teacher's punishment
136. Teacher's cultural promises
137. Teacher shares decisions
138. Teacher diagnoses and prescribes for math
139. Teacher's control by peer pressure
140. Teacher's personalization
141. Student eagerness
142. Reading specialist
143. Teacher uses aides for reading
144. Successful students known by peers
145. Good readers known by students
146. Identification of successful students by teacher
147. Student independent learning
148. Free mobility of students
149. Teacher recognition seeking behavior
150. Teacher's desk used for security
151. Cooperation of students
152. Restrictive seating arrangements (spatial rigidity)
153. Teacher politeness
154. Teacher's failure to correct student errors
155. teacher emphasis on discipline and control
156. Explaining manners (socialization)
157. Instructional activities predominating
158. Emphasis on housekeepping duties
159. Equality emphasized by teacher
160. Student(s) manipulate teacher
161. Students' freedom to laugh
162. Teacher enjoys teaching
163. Teacher evokes student participation
164. Student's attentiveness
165. Student's decoding teacher's intent
166. Teacher lack of responsiveness
167. Teacher's responsiveness
168. Role clarity
169. Consistency of message
170. Teacher modeling and student imitation
171. Confused students
172. Voting
173. Student's elation of own correct response
174. Instruction by manipulation of materials
175. Teacher uses aides for math
176. Continuity
177. Repetition or memory as instructional technique
178. Filling time
179. Teacher participates in routine tasks (equalizing)
180. Teacher as consultant
181. Student in charge of class
182. Cooperativeness between students
183. Seating structure discourages peer teaching
184. Teacher inability to tolerate children
185. Democratically run classroom
186. Appropriateness of tasks
187. Engagement
189. Spontaneity on the part of students
190. Clearly stated directions
191. Social closeness
192. Inappropriate expectations
193. Practice and recitation of material (continuity)
194. Contract teaching
195. Rigid spatial structure
196. Teacher inability to carry out lesson
197. Teacher threats follow-through inconsistent
198. Rush to finish
199. Putting heads down
200. Teacher's emphasis on orderliness
201. Student recognition seeking behavior
202. Teacher gives an example
203. Reading in unison
204. Teacher insecurity (II) vs. #126
205. Teacher recognition seeking (II)
206. Cooperativeness (II)
207. Student responsbility for self
208. Pupil orientation
209. Modeling (II)
210. Demanding proof
211. Students' open resistance

TABLE 4

61 Dimensions for Comparing Known Sample Classrooms

1. abruptness (T): unanticipated "switching" by teacher, e.g., from instruction to classroom management, to behavior management, to instruction, to behavior management.
2. accepting (T): teacher reacts constructively (overt, verbal, non-verbal) to students' feelings and attitudes
3. adult involvement (C): adults other than the teacher are allowed to instruct.
4. attending (T): teacher actively listens to what a student is saying, reading, reciting.
5. awareness of developmental levels (T): teacher is aware of a student's emotional, social educational needs and therefore assigns tasks appropriate for these.
6. being liked (T): teacher seeks approval from students in an ingratiating manner, often at expense of instruction.
7. belittling (T): teacher berates child in front of others.
8. competing (T): competition, outdoing others is emphasized by the teacher.
9. complimenting (control) (T): teacher's action reinforces student(s) whose behavior is in the right direction.
10. consistency of message (control) (T): teacher gives a direction or a threat and follows through with it.
11. conviviality (C): warmth, family-like quality to classroom interaction; good feelings between teacher-students, students-students.
12. cooperation (S): students cooperate with other students, teacher; willingness on part of students to help each other.
13. defending (T): teacher defends a student from verbal or physical assault by another.
14. defiance (S): a student's open resistence to teacher direction; refuses to comply.
15. democracy (T): teacher provides opportunities to involve students in decision-making re class standards, instruction, procedures, etc.
16. distrust (T): teacher expresses doubt for validity of student's work or behavior.
17. drilling (T): teacher emphasizes regularization, rote memory, retrieval of facts on part of student learning.
18. encouraging (T): teacher admonishes student effort in order to motivate them.
19. engagement (S): students express eagerness to participate, appear actively, productively involved in learning activities.
20. equity (T): teacher appears to divide her time, attention equally among all students.
21. ethnicity (T): teacher expresses positive, informative comments about racial, class, ethnic contributions; encourages class discussion about cultural contributions.
22. excluding (T): teacher banishes student from class activity--to corner, cloakroom, out of room, etc.
23. expectation (T): teacher attributes scholastic problems or predicts success for student on basis of past information or student's "background."
24. filling time (T): teacher fills "empty" time periods with "busy work".
25. flexibility (T): teacher adjusts instruction easilty to accommodate change in plans, time schedule, absenteeism, or change of students' behavior.
26. gendering (T): teacher assigns roles on basis of male or female (boy-girl) and reinforces these.
27. harrassing (T): teacher taunts, pesters, nags, hazes, "puts down," or physically hits a student.
28. ignoring (T): teacher appears to deliberately "not hear" or "not see" so as to treat a student as being invisible.
29. illogical statements (T): teacher makes a statement whose consequences would be ridiculous if carried out.
30. individualizing (T): teacher assigns to each student learning tasks designed to match his/her individual abilities and interests.

TABLE 4 (Continued)

31. job satisfaction (T): teacher seems to enjoy teaching.

32. knowledge of subject (T): teacher seems confident in teaching a given subject, and demonstrates a grasp of it.

33. manipulation (S): student is able to get on demand a desired response from the teacher.

34. mobility (S): students move freely and purposefully around the room; teacher allows students to work at places other than at their assigned seats.

35. mobility (T): teacher moves spontaneously about the room.

36. modeling/imitation (S): students copy teacher's behavior, and are encouraged to do so by teacher.

37. monitoring learning (T): teacher checks in on student's progress regularly and adjusts instruction accordingly.

38. moralizing (T): teacher emphasizes goodness vs. badness, verbally expresses ideal behavior model.

39. oneness (T): teacher treats whole group as a "one" often in order to maintain peer control.

40. openness (T): teacher verbally acknowledges to students feelings of anger or frustration, admits mistakes, expresses need for self-improvement.

41. open questioning (T): teacher asks questions which call for interpretive responses and are open-ended.

42. optimism (T): teacher expresses positive, pleasant, optimistic attitudes and feelings.

43. pacing (T): teacher appears to perceive learning rate of students and adjusts teaching pace accordingly.

44. peer teaching (S): students help other student instructionally and are encouraged to do so, whether "olders" with "youngers" or students of same age group.

45. personalizing (T): teacher calls on students by name.

46. policing (T): undue emphasis on quietness, orderliness, good behavior, and teacher spends disproportionate time with monitoring student behavior and controlling for discipline.

47. politeness (T): teacher requests rather than commands, uses "please" and "thank you", encourages same in student-student interaction.

48. praising (T): teacher verbally rewards student.

49. promoting self-sufficiency (T): teacher encourages students to take responsbility for their own classwork.

50. recognition-seeking (T): teacher calls attention to self for no apparent instructional purpose.

51. rushing (T): teacher does not give students adequate response time, or answers for them; is tied to a pre-set time limit, and hurries students to finish work.

52. sarcasm (T): teacher responds in a demeaning manner, uses destructive/cutting remarks.

53. shaming (T): teacher instills guilt in students for their behavior in order to establish control.

54. signaling (control) (T): teacher uses body language, non-verbal signals to change students' behavior.

55. spontaneity (T): teacher capitalizes instructionally on unexpected incidents that arise during class time.

56. stereotyping (T): teacher labels and judges students by socio-economic, ethnic, or racial characteristics.

57. structuring (T): teacher prepares students for lesson by reviewing, outlining, explaining objectives, summarizing.

58. teacher made materials (T): teacher provides instructional materials other than textbooks, and arranges for their use by students.

59. time fixedness (T): teacher emphasizes promptness, begins and ends activities by clock rather than by student interest.

60. waiting (T): after asking a question, teacher waits in silence for student responses or waits in silence after student response before reacting.

61. warmth (T): teacher seeks contact with students, talks with them shows affection toward them.

TABLE 5

Example from Raters Manual

ABRUPTNESS (T)

We mean:

-- teacher changes from one activity or lesson to another without advising students
-- unanticipated change without "tying up" what was in progress
-- pupils surprised or confused by teacher's change in behavior
-- teacher switches from instruction to behavior management and back again

Examples:

1. Teacher, "Okay, get out your readers and turn to page 195." Students, "But we haven't finished our math worksheet." Teacher, "Sorry, maybe you can find time to do it later."

2. During reading lesson, teacher is listening to students in a small group reading aloud. Several times she interrupts whoever is reading in order to administer discipline to someone in another part of the room.

3. Students have been working diligently, but noisily. After several warnings, teacher says, "All right. Put your books away. Since you already know the materials, it's quiz time."

We do not mean:

-- smooth transition between work periods
-- eases in from one activity to the next

Examples:

1. Teacher makes sure that students understand what they are to do before starting the activity.

2. Teacher systematically monitors work, gives assistance when or before it is needed.

3. Before math begins students may get a drink and relax for a moment. Then they get their math books and papers and begin their work.

TABLE 6

Paired Comparisons for 61 Dimensions

Dimension	SECOND GRADE MATHEMATICS			SECOND GRADE READING			FIFTH GRADE MATHEMATICS			FIFTH GRADE READING			% Rater Agreement
	More Effective Teachers*	Less Effective Teachers*	p	More Effective Teachers*	Less Effective Teachers*	p	More Effective Teachers*	Less Effective Teachers*	p	More Effective Teachers*	Less Effective Teachers*	p	
(-) 1. abruptness (T)	13	23	.0475	27	9	.0013	11	25	.0099	10	26	.0035	.81
(+) 2. accepting (T)	26	10	.0035	25	11	.0099	24	12	.0228	26	10	.0035	1.00
(+) 3. adult involvement (C)	28	8	.0003	23	13	.0475	23	13	.0475	23	13	.0475	.75
(+) 4. attending (T)	26	10	.0035	30	6	.00003	25	11	.0099	25	11	.0099	.81
(+) 5. awareness of developmental levels (T)	23	13	.0475	29	7	.0002	22	14	.0918	26	10	.0035	.88
(-) 6. being liked (T)	11	25	.0099	15	21	.1587	9	27	.0013	14	22	.0918	.69
(-) 7. belittling (T)	11	25	.0099	12	24	.0228	12	24	.0228	12	24	.0228	.81
(+) 8. competing (T)	19	17	.7414	12	24	.0456	14	22	.1836	14	22	.1836	.81
(+) 9. complementing (control) (T)	20	16	.2514	26	10	.0035	23	13	.0475	19	17	.3707	.75
(+) 10. consistency of message (control) (T)	26	10	.0035	24	12	.0228	28	8	.0003	27	9	.0013	.69
(+) 11. conviviality (C)	25	11	.0035	30	6	.00003	23	13	.0475	27	9	.0013	.75
(+) 12. cooperation (S)	26	10	.0035	29	7	.0002	24	12	.0228	27	9	.0013	.81
(+) 13. defending (T)	19	17	.3707	21	15	.1587	17	19	.3707	23	13	.0475	.63
(-) 14. defiance (S)	11	25	.0099	4	32	.000003	13	23	.0475	7	29	.0002	.81
(+) 15. democracy (T)	25	11	.0099	23	13	.0475	22	14	.0918	26	10	.0035	.75
(-) 16. distrust (T)	10	26	.0035	12	24	.0228	14	22	.0918	17	19	.3707	.88
(+) 17. drilling (T)	14	22	.1836	13	23	.0950	19	17	.7414	12	24	.0456	.69
(+) 18. encouraging (T)	28	8	.0003	30	6	.00003	21	15	.1587	25	11	.0099	.75
(+) 19. engagement (S)	27	9	.0013	28	8	.0003	23	13	.0475	30	6	.00003	.88
(+) 20. equity (T)	23	13	.0475	27	9	.0013	22	14	.0918	26	10	.0035	.94
(+) 21. ethnicity (T)	22	14	.0918	26	10	.0035	18	18	.5000	25	11	.0099	.69
(-) 22. excluding (T)	11	25	.0099	16	20	.2514	13	23	.0475	13	23	.0475	.75
(-) 23. expectation (T)	11	25	.0099	13	23	.0475	14	22	.0918	14	22	.0918	.75
(-) 24. filling time (T)	11	25	.0099	9	27	.0013	10	26	.0035	9	27	.0013	.81
(+) 25. flexibility (T)	27	9	.0013	27	9	.0013	22	14	.0918	23	13	.0475	.88
(+) 26. gendering (T)	17	19	.7414	11	25	.0198	16	20	.5028	16	20	.5028	.63
(-) 27. harrassing (T)	12	24	.0228	8	28	.0003	14	22	.0918	10	26	.0035	.75
(-) 28. ignoring (T)	10	26	.0035	7	29	.0002	16	20	.2514	9	27	.0013	.75
(-) 29. illogical statements (T)	13	23	.0475	8	28	.0003	12	24	.0228	11	25	.0099	.50
(+) 30. individualizing (T)	26	10	.0035	28	8	.0003	20	16	.2514	23	13	.0475	.94
(+) 31. job satisfaction (T)	29	7	.0002	29	7	.0002	22	14	.0918	28	8	.0003	.75
(+) 32. knowledge of subject (T)	27	9	.0013	29	7	.0002	25	11	.0099	28	8	.0003	.88
(-) 33. manipulation (S)	17	19	.3707	6	30	.00003	9	27	.0013	8	28	.0003	.69
(+) 34. mobility (S)	27	9	.0013	27	9	.0013	22	14	.0918	21	15	.01587	.69
(+) 35. mobility (T)	26	10	.0035	28	8	.0003	19	17	.3707	17	19	.3707	.69
(+) 36. modeling/imitation (S)	22	14	.0918	23	13	.0475	19	17	.3707	21	15	.1587	.63
(+) 37. monitoring learning (T)	29	7	.0002	29	7	.0002	23	13	.0475	24	12	.0228	.88
(-) 38. moralizing (T)	7	29	.0002	8	28	.0003	20	16	.2514	21	15	.1587	.63
(-) 39. oneness (T)	11	25	.0099	12	24	.0228	12	24	.0228	11	25	.0099	.63
(+) 40. openness (T)	20	16	.2514	26	10	.0035	22	14	.0918	21	15	.1587	.56

TABLE 6 (Continued)

Dimension	SECOND GRADE MATHEMATICS			SECOND GRADE READING			FIFTH GRADE MATHEMATICS			FIFTH GRADE READING			% Rater Agreement
	More Effective Teachers*	Less Effective Teachers*	p	More Effective Teachers*	Less Effective Teachers*	p	More Effective Teachers*	Less Effective Teachers*	p	More Effective Teachers*	Less Effective Teachers*	p	
(+) 41. open questioning (T)	24	12	.0456	27	9	.0026	26	10	.0070	19	17	.7414	.63
(+) 42. optimism (T)	28	8	.0003	29	7	.0002	23	13	.0475	27	9	.0013	.88
(+) 43. pacing (T)	27	9	.0013	26	10	.0035	23	13	.0475	25	11	.0099	.69
(+) 44. peer teaching (S)	26	10	.0035	19	17	.3707	18	18	.5000	24	12	.0228	.94
(+) 45. personalizing (T)	24	12	.0228	25	11	.0099	23	13	.0475	22	14	.0918	.56
(-) 46. policing (T)	7	29	.0002	9	27	.0013	14	22	.0918	14	22	.0918	.81
(+) 47. politeness (T)	23	13	.0475	26	10	.0035	25	11	.0099	21	15	.1587	.88
(+) 48. praising (T)	29	7	.0002	29	7	.0002	21	15	.1587	23	13	.0475	.81
(+) 49. promoting self-sufficiency (T)	25	11	.0099	30	6	.00003	23	13	.0475	24	12	.0228	.44
(-) 50. recognition seeking (T)	10	26	.0035	11	25	.0099	12	24	.0228	12	24	.0228	.69
(-) 51. rushing (T)	10	26	.0035	12	24	.0228	16	20	.2514	18	18	.5000	.75
(-) 52. sarcasm (T)	10	26	.0035	12	24	.0228	11	25	.0099	18	18	.5000	.75
(-) 53. shaming (T)	13	23	.0475	8	28	.0003	12	24	.0228	14	22	.0918	.63
(+) 54. signaling (control) (T)	17	19	.7414	15	21	.3174	8	28	.0006	8	28	.0006	.69
(+) 55. spontaneity (T)	26	10	.0035	28	8	.0003	25	11	.0099	23	13	.0475	.81
(-) 56. stereotyping (T)	16	20	.2514	11	25	.0099	14	22	.0918	16	20	.2514	.69
(+) 57. structuring (T)	24	12	.0228	28	8	.0003	26	10	.0035	25	11	.0099	.75
(+) 58. teacher made materials (T)	21	15	.1587	29	7	.0002	21	15	.1587	19	17	.3707	.75
(-) 59. time fixedness (T)	11	25	.0099	10	26	.0035	18	18	.5000	18	18	.5000	.81
(+) 60. waiting (T)	26	10	.0035	27	9	.0013	22	14	.0918	23	12	.0475	.75
(+) 61. warmth (T)	27	9	.0013	26	10	.0035	22	14	.0918	24	12	.0228	.81

* Number of times in 36 paired comparisons the more effective/less effective teacher's protocol was rated as evidencing more of the dimension.

Protocol Number: 06
Name of Researcher:Gail
Date of Observation:
Subject of Observation: 2nd Grade Class, Open Class-

1. room, with two team teacher and two other adults, this
2. is a joint observation with Elizabeth. I will be
3. observing two reading groups today, simultaneously,
4. including 9 children. Out of the nine children, 2 are
5. girls, 7 are boys.
6.

8:30 Noise level 2

7. At 8:30 the noise level is 2. The children have just been
8. let into the classroom, taking their coats off and
9. wandering around the room. Several boys are in the corner
10. fighting, and some girls are sitting on the floor
11. playing a puzzle. The teacher is walking back and forth
12. in the back of the classroom not attending the children.
13. The noise continues and the children are running
14. around. There is much confusion in the room. Two teachers

8:35

15. stand at the desk talking to one another. At 8:35,
16. Mrs. Tyler leaves the room. The team teacher
17. stays seated behind the classroom at her desk. At 8:40
18. Mrs. Tyler comes back into the room. She walks to the
19. desk at the far left hand side of the classroom,
20. which is a round table, and sits on the edge. She says
21. "Blue Group, get your folders and go up in the front.
22. Green Group, come here." Noise level drops to 1, and
23. the children begin to follow her orders. She says,
24. "Anybody loose a quarter.." No one responds, and she
25. repeats the question again with irritation in her voice.
26. She says I know someone found , someone lost a quarter
27. because it was found in the coat room. Look in your
28. pockets and see." No one says anything. She now
29. stands up and pulls a pile of workbooks from across the
30. table over to her. They are the _____ reading work-
31. books. She opens one of them on the top and says,
32. "Ah Daniel!" She says this with a loud sharp voice.
33. She continues, "Your work yesterday was not too bad
34. but you need some work. Evidently there are still some
35. words you don't understand." She thumbs through the rest
36. of his lesson. Danny is standing at the outside of
37. the circle around her, not listening to what she is saying.
38. Mrs. Tyler now stands and gives instructions to the Green
39. Group. She tells them to go through 8 through 13, reading
40. the two stories between those pages and to go over the
41. work in the workoooks that she is about to give back.
42. She tells them that they may seat any place but
43. not together and she says, "And I don't want any funny
44. business." She now opens the next workbook which is
45. Nicolle's. She tells Nicole that she is having the
46. same problem that Danny is having without specifying
47. further. Nicole looks up at her with an expectant look
48. on her face. She then looks at a third book and says
49. Michelle you're having the same problem. She says,
50. "Snatch means to grab. Beach, what does it mean? Michelle
51. doesn't answer. She has her finger in her mouth and looks
52. anxious. The teacher closes the workbook and pushes it
53. to Michelle. Michelle takes it and walks away, with
54. Nicole. Teacher then opens the next workbook and says,
55. Mike, I don't appreciate all these circles. She points

Figure 1. A sample protocol.

We are looking for THINGS THAT DISCRIMINATE BETWEEN CLASSROOMS. When you find one DESCRIBE IT:

A. What it is and/or looks like (descriptors, characteristics, connotations, synonyms)

B. What it is not (antonyms, descriptors, characteristics, non-examples)

GIVE EXAMPLES:

1.

2.

3.

LABEL IT: _____

Figure 2. Teacher, Student, or Instructional Characteristics Card for Generating Dimensions.

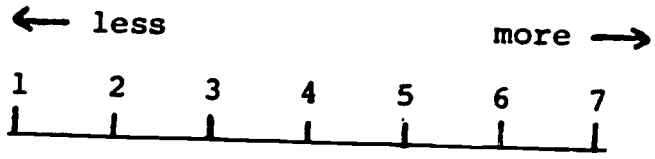
Name of Rater: _____

Date: _____

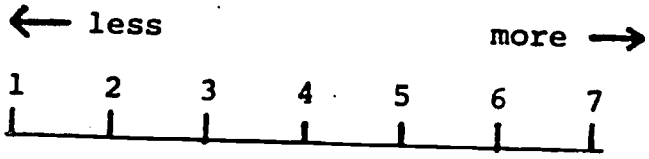
Classrooms being compared: _____ x _____

Grade Level: _____ Subject: _____

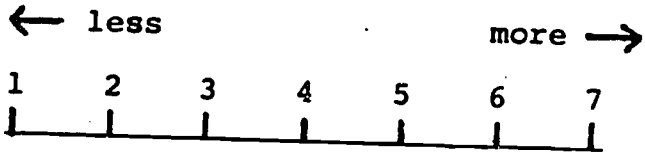
1. Abruptness (T)



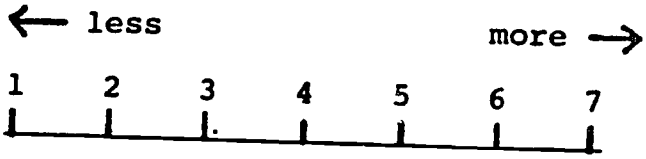
2. Accepting (T)



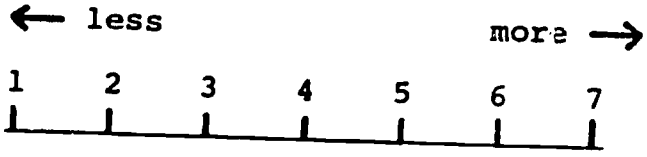
3. Adult Involvement (C)



4. Attending (T)



5. Awareness of developmental levels (T)



6. Being liked (T)

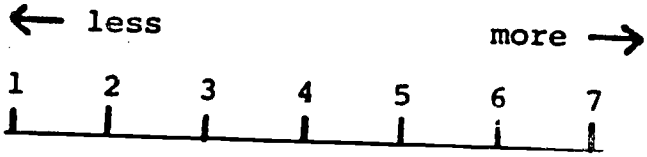


Figure 3. Classroom Comparison Instrument on 61 Dimensions.

Less Effective Teachers \ More Effective Teachers	2	4	11	14	12	8	7	20	5	13
15	K				S	M				R
18		K		S			M		R	
10			K S					N Q		
16		R		K			Q		N	
1	R				L	Q				N
3	N				Q	L				T
6		O		P			L		T	
17			O P					I T		
19		P		O			T		M	
9	P				O	S				M

Capital letters in cells (K, R, etc.) represent raters assigned to protocol pairs.

Figure 4. Second Grade Reading: Paired Comparisons of Known Sample Teachers

END

DEPT. OF HEW

NAT'L INSTITUTE OF EDUCATION

ERIC

DATE FILMED

JAN 3 77