One of the main advantages of the naturalistic approach to classroom research is that the observations are of natural phenomena, occurring in circumstances which the researcher ultimately hopes to generalize. Some problem areas are associated with this approach, however. Observers must obtain complete and accurate records of classroom activities and teacher and student behavior. They must also maintain objectivity in their narrative records. The difficulty also exists that the presence of the observer may intrude or interfere with normal classroom functioning. This paper presents some of the considerations made and techniques used to select and train observers for a major naturalistic classroom observation study, the Junior High Classroom Organization Study. Priorities used for selecting observers are described, including classroom experience, either as a teacher or observer, and writing skills. A description is given of the activities which were engaged in during the training week. It is reported that each of the three problem areas that are associated with naturalistic studies was dealt with and some success was achieved in ameliorating them: It is pointed out that it is necessary to choose observers carefully, train them extensively, and supervise them closely. (JD)
Developing an Effective Research Team for Classroom Observation

Barbara S. Clements
Carolyn M. Evertson

Research and Development Center for Teacher Education
The University of Texas at Austin


(R&D Report No. 6103)

Research and Development Center for Teacher Education
The University of Texas at Austin
February, 1980

This study was supported in part by the National Institute of Education under Contract OB-NIE-G-80-0116, The Classroom Organization and Effective Teaching Project, The Research and Development Center for Teacher Education, The University of Texas at Austin. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by that office should be inferred. Requests for reprints should be addressed to Communication Services, R&D Center for Teacher Education, Education Annex 3.203, The University of Texas at Austin, Austin, Texas 78712.
Developing an Effective Research Plan for Classroom Observation

A number of classroom research studies have used detailed anecdotal records taken by observers as primary data sources, the global-descriptive or naturalistic approach. This methodology "focuses on the development of hypotheses rather than the testing of them" (Tikunoff, Berliner, & Rist, Note 1, p. 38). Certain assumptions are drawn from theories or past research and are used as a focus of attention in classroom observations. After data are collected, these assumptions are refined to reflect empirical fact. One of the main advantages of the naturalistic approach is that the observations are of natural phenomena, occurring in the "circumstances to which the researcher ultimately hopes to generalize" (Tunnel, 1977, p. 432). The concern with these studies is different from so-called quantitative classroom studies, using ratings or counts of behavior. Instead of an emphasis on reliability of observers in making the ratings, the emphasis is on the ability of observers to faithfully record events and activities in the classroom.

A number of issues and problems associated with the naturalistic approach have been described in detail by Tikunoff, Berliner, and Rist (Note 1), Kounin (1970), Wolcott (1975), and Tunnel (1977). Among them are:

1. Observers must obtain complete and accurate records of classroom activities and teacher and student behaviors. Observers tend to selectively notice and record only very impressive events, such as misbehavior by a particular student, using his/her own frame of reference or what is impressive to him/her. A complete narrative record
should contain a room description and details as to what parts of the room are included in class activities. Activities should be described in detail, stating what the teacher and students are doing. Copies of handouts and sketches or descriptions of other materials used should be included. Actual quotes, where possible, are extremely valuable. Therefore, good writing skills, high fluency, and an eye for detail were seen as necessary observer characteristics.

2. Observers must maintain objectivity in their narrative records. There is a propensity for observers to include nonobjective or nondescriptive entries such as labels, summaries and evaluations. Observers have usually had intimate contact with the classroom setting since the age of six and may have preconceived notions as to what the ideal situation should be. Observers also may identify or sympathize with particular students or the teacher in certain situations. Observers should try to remain aloof from the affairs of the classroom and record events objectively.

3. Because there is an adult (the observer) present in the classroom who is not normally there, there is a possibility of intrusion upon or interference with the normal classroom functioning and interactions. In order to provide a faithful rendering of a natural classroom, it is necessary for an observer to "blend into the walls," such that the teacher and students have received the same experience they would have received if the observer was not there.

Although it is impossible to control completely these potential problems, their effects can be minimized by careful selection, training, and supervision of observers. This paper presents some of the
considerations made and techniques used to select and train observers for a major naturalistic classroom observation study.

Methods and Procedures

The Junior High Classroom Organization Study (Evertson, Emmer, & Clements, Note 2), done in 1978-79, drew much of its methodology and instrumentation from the Classroom Organization Study (Evertson, Anderson, Emmer, & Clements, Note 3), done in elementary school classrooms the year before. Twenty-five English and 26 math teachers in 11 junior high schools participated in the Junior High Classroom Organization Study. Two classes of each teacher were observed. One class of each teacher was seen on the first day of school and each teacher was observed an additional 8-10 times in both classes during the first three weeks of school. Nineteen observers were used during the first three weeks of school with each teacher being seen by two observers. Thereafter, five observers continued to observe approximately 10 teachers each. The typical observation lasted for 55 minutes. An attempt was made to observe each teacher's classes every three weeks and on different days of the week during the rest of the school year.

Selection of Observers

There were several considerations made in the selection of observers for the Junior High Classroom Organization Study. One of the prerequisites for selection was classroom experience, either as a teacher or observer. It was hoped that by obtaining observers with such experience, several important criteria would be met. Observers with classroom experience have experienced the classroom in both the role of the student and that of a teacher or observer. Therefore, it was
believed that they would be more objective in reporting classroom events, behavior, and interactions. In addition, observers with classroom experience have more knowledge of the setting, activities, problems, and constraints and would therefore be better able to give accurate and detailed reports. Finally, it was believed that teachers would be more willing to participate and comfortable with our observers present, knowing that they had teaching or other classroom experience.

Another prerequisite for selection was writing skill. Observers need to be able to relate events factually and sequentially and write coherently. The majority of the observers used in our study were graduate students from the departments of Educational Psychology or Curriculum and Instruction at The University of Texas at Austin who were beyond their first year of graduate school. It was felt that advanced graduate students would have adequate writing skills and be able to record the detail we desired. Other observers were R&D Center staff members and former teachers.

An additional prerequisite was availability. Classroom observations are very time-consuming and tiring. In our study, for every hour spent in the classroom, an additional 1-1/2 to 2 hours were spent out of the classroom, organizing notes, taping the narrative record, and filling out forms. During the first week of school, our observers worked full time, each day spending three hours in the classroom, approximately 4-1/2 hours outside of the classroom, and 1/2 hour in transit. During the second and third weeks of school, our observers worked half-time, spending from seven to nine hours in the classroom each week, at varying times of the day.
Observer Training

Training for observers took place for a week prior to the first day of school. Observers met with R&D Center staff members for approximately 25 hours of training, or five hours per day. Observers were asked to study and do some work at home in addition to the work during the training sessions. A schedule of training was developed to give adequate time to discuss each rating form and narrative writing and practice each one, both individually and along with other parts of the study. A brief description of the schedule follows.

Monday. After introduction of the staff members and observers, an overview of the Junior High Classroom Organization Study was given, including a brief description of the Classroom Organization Study, the preliminary results of that study, and the purpose of the new study. Observers were given copies of the Observer Notebook which contained sufficient materials for training and the classroom observations. The notebooks had a section for each observation procedure, including the Narrative Record, the Student Engagement Ratings (SERs), the Component Ratings, and the Time Log. An additional section was included for miscellaneous information such as directions to the schools, lists of codes, etc. Following the introduction to the study was a detailed discussion of the Student Engagement Ratings. These measures, done every 15 minutes during the class period, were counts of on- and off-task behaviors. Videotapes of junior high or upper-elementary school classrooms were used for practice. During the afternoon, noting time intervals on the narrative forms and filling out time log forms were discussed and practiced using videotapes. Additional practice of the SER was done at this time, also.
Tuesday. Feedback was given to observers as to their performance on the exercises of the previous day. The rest of the meetings on this day centered on narrative writing. Examples of good narratives from the third-grade study were presented and general guidelines and areas of emphasis for the junior high study were discussed. In addition, a narrative done in a junior high school classroom was presented as an example of the types of activities and behaviors that could be expected in this study. Extensive videotape practice of narrative writing was done, with some immediate feedback as to what should have been noted and described. Observers turned in expanded narrative notes to be checked by staff members overnight.

Wednesday. Feedback was given to observers concerning their narrative notes from the previous day and good examples were read to the group. Some time was spent clarifying areas of confusion and emphasizing areas of importance in the study. Additional videotape practice was done using the narrative record, Student Engagement Ratings and Time Logs. Materials were again handed in for checking by staff members.

Thursday. Observers were returned their materials from the previous day with general and specific feedback. Component Ratings, a set of general questions about teacher and class characteristics, were discussed and practice done using videotapes. This was followed by additional discussion to clarify areas of confusion. In the afternoon a final 30-minute videotape was viewed and observers were instructed to complete all forms and write an expanded narrative record. These forms were turned in and checked by staff members overnight.
Friday. Feedback was given to observers about their narrative records and forms, and areas of remaining confusion discussed. General guidelines for observers were discussed. These included how to dress to be unobtrusive and professional, what could be said to teachers, how to avoid contacts with students, when to arrive for class, and what to do if the teacher left the room. Additional information about dictating narratives and handing in materials was discussed, followed by the distribution of observation schedules, additional forms, tape recorders, and tapes.

Observers were encouraged to read the guidelines at home and ask questions when needed. At all times during the week of training and the weeks of observation, there were staff members available to answer questions. Specific feedback on use of forms was given to each observer daily. Observers were receptive to the feedback and all appeared to understand the amount of detail and accuracy desired.

Reliability During the First Three Weeks

Observer materials were checked periodically during the first three weeks by R&D staff members for level of detail and correct completion of the forms. Questionable ratings were discovered by comparing ratings sheets filled out by the two observers for a particular class. Observer agreement was computed at the end of the first three weeks on the Student Engagement Ratings and Component Ratings and moderate to high agreement was found. In addition, at least moderate stability across classes was found.

Additional Observer Meetings

An additional observer meeting was held on Friday afternoon of the
second week of school to give observers an additional 10 component ratings to fill out during the third week of school.

On the last day of the third week of school, observers met with R&D staff members to discuss a set of Summary Component Ratings to be filled out and some questions to answer on the teachers they had observed. Observers were given two weeks to complete the observation forms from the first three weeks of school and the summary forms.

Reliability Throughout the Rest of the Year

Regular observer meetings for the five continuing observers were held approximately once every three weeks throughout the rest of the school year. These meetings served primarily to clarify areas of confusion and to give feedback to observers on their performance.

During the three week cycles prior to Christmas, each observer saw two teachers normally seen by another observer and completed regular observation. The Component Ratings and Student Engagement Ratings were then checked against the Summary Component Ratings and a sampling of Student Engagement Ratings from the first three weeks. It was expected that student engagement ratings would not significantly vary from the first three weeks. Allowances were made for differences in activities in checking these ratings. On the Component Ratings, a check was made to see which dimensions had excessive discrepancies from the summary ratings from the first three weeks, if the two previous observers agreed. These dimensions were then checked against the written observation to see if they were consistent with what was written. These dimensions were also discussed with the observer and attempts made to be sure that all observers understood the ratings and were rating them consistently.
After Christmas a new reliability system was instituted. During each cycle, each observer saw a teacher with the regularly scheduled observer. Arrangements were made with the teachers in advance so that the observers could sit beside each other and have the same vantage point for taking Student Engagement Ratings. The observers would agree at what time to do the ratings but they would not consult with each other while doing them. In addition to the narrative and Student Engagement Ratings, observers also did the Component Ratings and Time Logs. The Student Engagement Ratings were then compared to see if there was observer agreement. The Component Ratings were checked against each other and differences of more than one point on the scale were discussed with the observers and checked against the narratives. Because good observer agreement was obtained in these checks, changes were not made on any of these forms. Discussions were held in the regularly scheduled observer meetings to maintain reliability.

Spot checks were also done on the narratives to be sure that observers were providing the desired detail and coverage of all major points.

Coping With Problems

Much was learned about selecting and training observers during this study. Each of the three problem areas discussed that are associated with naturalistic studies was dealt with and some success was achieved.

It was found that a high level of complete and accurate detail on narrative records could be expected of and received from observers. This was obtained partially through extensive training and feedback and frequent reliability checks. In general, observers were able to record ratings every 15 minutes during the observation and without a
significant reduction in detail. In particular, it was found that richer and more complete detail could be obtained by having observers take whatever notes could be recorded during the observation and then tape-record the narrative soon after the observation, adding additional information recalled but not recorded in their notes. Observers quickly became adept at using abbreviations and signals in their notes to cue their memories when later using the tape recorder. In addition, observers, using the set of guideline questions, were encouraged to record comments or additional detail at the end of the narrative.

The issue of objectivity was dealt with in two ways. Observers with classroom experience were chosen with the expectation that they would be able to more objectively record classroom events, having had experience both as students and as teachers or observers. In addition, observers were allowed to express parenthetically, or at the end of the narrative, their opinions and attitudes about the observations. As a result, narrative readers were better able to separate actual classroom events from observer opinions and allow for any biases of the observer.

It is not possible to accurately assess the impact of our observers on the classes in the study. For instance, it was impossible to determine whether teachers presented specially prepared lessons on the days they were observed. Lessons and activities did appear to be consistent through the year for participating teachers. Many teachers reported that they usually forgot when the observer would be there until he/she entered their rooms. Teachers were asked in a questionnaire, filled out at the end of the school year, if they thought the observer had an effect on them or their students. At the beginning of school, it was suggested that teachers tell the students who the observer was and
what he or she was doing. Most teachers reported that after the introduction was made, both teacher and students forgot the observer was present. Several teachers mentioned that because the observer came so often, he/she was familiar and the students became used to having him/her there. Other than one class's apparent reluctance to respond, some curiosity as to what was being written and occasional showing off were the only reported effects of the observers' presence. Many teachers expressed surprise at how little the effect was on student behavior.

Of the eight teachers who responded that their teaching had been affected, one responded that it had made him "meaner" and one responded that she was stiffer in her presentations to the class. The other six teachers responded that the changes in their behavior were positive. One teacher reported he was "less harsh" with his students, another became more aware of her school year as a whole. The other teachers reported they tried a little harder to be good.

A number of teachers' comments confirmed our belief that our observers had developed good rapport with their teachers. Teachers complimented the observers, stating that they were unobtrusive, professional, and supportive of the teacher. One teacher mentioned that the reason some of her fears of having an observer were allayed was, "After I found that she was a classroom teacher, I felt the observations would be objectively recorded and noted with more empathy than those of my immediate administrators."

Observer comments supported the teacher reports. Observers reported they had little or no effect on student behavior in the classroom and that teachers seemed comfortable having them. End-of-year
interviews of the teachers were conducted by the observers and the information obtained confirmed the friendliness and rapport of observers and teachers.

While the quality of the observational data was very high, a number of problems arose that were not adequately dealt with and will require additional attention in future studies.

One of the major problems discovered during observer monitoring and narrative reading was the short length of some of the narrative records. It was found that a few observers were not recording the level of desired detail, a problem not evidenced during training sessions. Several things can be done to deal with this problem. It would be advisable to pay particular attention to verbal skills during the interview of potential observers, in an attempt to discover the person's probable capacity for relating details. Testing of short-term memory or hand-eye coordination could be done to supplement interview information in the selection of a team of observers. In addition, it would be preferable to hire more than the needed amount of observers and set a criterion to which observers must be trained in order to continue in the study. The advantages of training to a criterion would be more detailed, complete narrative records and better observer agreement on variables. We plan, in future studies, to require additional practice with tape recording of narrative records and to give more specific feedback regarding what could have been noted and what each observer did note during practice sessions.

An additional problem regarding the Narrative Records was the tendency of some observers to give global assessments of classroom activities. General phrases such as "acting bad or acting good" which
are not elaborated give very little specific information about what was actually happening in the classroom. While our observers were given a long list of specific descriptive terms that could be used, we will probably spend more time on the kinds of phrases and terms which are not specific enough. Additional stress on direct quotes, precise descriptions, and notation of time will be made.

In summary, it is necessary to choose observers carefully, train them extensively, and supervise them closely in order to receive complete, accurate, and objective records of classroom activities and teacher and student behaviors in a naturalistic classroom study.
Reference Notes


References

