Based on research using participant and nonparticipant observation in a natural setting, a study examined how teaching practices, classroom conditions, and student differences influenced the design of writing lessons in 14 English classes in five inner city schools. Data were collected from five sources--interviews, writing samples, questionnaires, school records, and over 300 full-period classroom observations. Results indicated that writing episodes in all classes averaged 16 minutes and consumed one-third of the total class time. Five kinds of assignments were observed: those based on reading, those based on ideas, skill assignments, rhetorical assignments that provided an audience emphasis or an explicit form, and multiple focus assignments that emphasized more than one of the other four types. In all classes, skill assignments were the most frequently made, and idea assignments took up the largest proportion of total class time. Teachers, however, tended to vary assignments to fit the needs of particular students. (FL)
THE VARIABLES IN LESSON DESIGN

One of the ideals in ethnographic research is to report the methodology in such detail that other researchers can use the report as an operating manual by which to replicate the study (Compte and Goetz, 1982:40). But journal articles regularly fail to report adequately their research design and methodology because journals require a brevity made possible by a common language for research techniques. But the absence of a technical shorthand for ethnographic research techniques means that journal articles, in order to be brief, must be incomplete (Pelto and Pelto, 1978). This article, like others not a complete guide for replication, gives secondary attention to results and primary attention to design and methodology, particularly formulating the problem and establishing validity, and reliability.

Formulation of the Problem

In the present study, the problem was shaped by two circumstances: first, the absence of observational studies of the teaching of writing in general and of inner city writing classrooms in particular; and second, the presence of conflicting notions about how to define teaching effectiveness. Until recently observational studies of teaching have focused primarily on the teaching of reading and mathematics and ignored writing (Brophy and Evertson, 1974). In 1980 and 1981 three separate groups of researchers completed studies of

The present study narrowed the focus found in Applebee’s study in two ways. First this present study examined the practices of teachers who emphasize writing, not all teachers. For his purposes, Applebee observed teachers who were willing to participate and presumably some or many of them did not regard the teaching of writing as one of the primary goals of their instruction. Second, this present study examined the teaching of writing in inner city schools, and these schools differ markedly from the city school used in Applebee’s study. One difference is that Applebee’s city school had a minority population of 20.7%, and the schools used in this present study had a minority population of 90% or more. The present study examined inner city schools because the most pressing problems confronting public schools today are in inner cities, and most experts agree that the ultimate test of any program of educational change is whether it can succeed there (U.S. Department of Health, Education, and Welfare, 1979:8).

In addition to the absence of observational studies in writing classes in inner city high schools, the present study was shaped by a second circumstance -- the presence in school districts of conflicting notions about how to define effective teaching. Teachers in one school district were being told "too much written assignment time" and "too much working with one student at a time" were negatively associated with effective teaching and that "drill and practice" and "short quizzes" were positively associated with effective teaching. (The Stallings Observation Strategy, Morgan Hill Unified, 1982).
Stallings based her ideal model of teaching on her study of correlations between instructional activities and CTBS scores (Stallings, 1980). A similar model of effective teaching, also adopted by some school districts, was what Barak Rosenshine called "direct instruction": "large groups, decision making by the teacher, limited choice of materials and activities by the students, orderliness, factual questions, limited exploration of ideas, drill, and high percentages of correct answers" (Rosenshine, 1979:47).

At the same time that direct instruction and the Stallings model were being proposed as models of effective teaching for all teachers of academic subjects, Charles Cooper was proposing a quite different model of the effective teaching of writing (Cooper, 1981: xi-xiii). Cooper emphasized the positive benefits of teachers working individually with students on their writing, increasing the writing of at least paragraph length and decreasing short answer quizzes, and giving substantial writing time in class when teachers can give help, all matters which Stallings and Rosenshine found to be negatively associated with effective teaching.

Why this difference in models of effective teaching? First, direct instruction and the Stallings' model may not apply equally well to all subjects. The Stallings' model, for instance, used CTBS scores, not scores on writing samples, as a measure of effectiveness. Rosenshine's direct instruction model was based on measures of gains in reading and mathematics, not writing. Students who receive "direct instruction" end up doing better on achievement tests, but on problem solving tests these students do worse than students who receive "open teaching" (Peterson, 1981:63). Because writing is a problem solving activity,
writing may require a different kind of instruction. The first question is, then, how does writing instruction look in classrooms taught by teachers who claim to be teaching writing?

A second explanation of the difference between Cooper’s model and that of Stallings and Rosenshine is that Cooper’s model appears to be based primarily on studies of how students learn to write in contexts outside the classroom, not how students learn or how writing is taught in classrooms. Studies from Emig (1971), Perl (1979), Halliday (1977) are examples of the types of studies that Cooper and others frequently use as a basis for recommending particular practices, but none of these studies observed students learning to write in classrooms. Both Emig and Perl were in schools, but their observations of student writers were outside of classrooms. Graves observed students learning to write in classrooms, but he narrowed his focus to a few students, leaving unreported what other students were doing and what strategies were used by the teacher (Graves, 1975). Cooper appears to discount the influence of classroom context on instruction, arguing that his recommended lessons do not appear in the classes observed by Applebee because of "lack of writing and poor instruction" (Cooper, 1981:xiii).

The question is whether classroom conditions influence the way writing is taught and whether these conditions might require teaching strategies quite different from those recommended by studies of learning outside of classrooms. Dreeben (1978), for one, has argued that "teachers deal with classroom conditions as much as they deal with individual students." Thus, studies of how individual students learn may not be an adequate source of information about how successful teachers teach. In other words, although the methods of direct instruction may not be the best way to learn to write in
general, some of the methods of direct instruction may be necessary to handle classroom conditions and, thus, may be the best way to learn to write in classrooms. To study the impact of conditions on teaching, one needs classrooms in which many different conditions are evident. Most scholars recognize that the problems of classroom conditions are more intense in inner-city classrooms than in other types of classrooms:

"To be sure, the kinds of problems noted are not confined to urban or inner-city secondary schools. In varying degrees such problems are found in secondary schools in all geographic locations. The degree of intensity of many school problems in suburban and rural areas, however, is thought to be significantly less than that found in inner-city schools" (Reed, 1982:7-8).

Barr and Dreeben have commented on the failure of many studies of teaching to consider the many variables that influence how lessons are designed by teachers:

"Two characteristics of classroom research limit its usefulness: the tendency to examine only a portion of the events occurring within classes, ignoring others which may bear on the productive process, and the tendency to ignore the contextual properties of classrooms" (Barr and Dreeben, 1980:151).

The first question --How is writing taught in writing classes in inner-schools?-- was then modified by a consideration of classroom and student variables. As a result, the central question of the study became 'How do classroom conditions and student differences appear to influence the way writing lessons are designed in inner-city secondary schools?' This question required an ethnographic research design using participant and non-participant observation in a natural setting.
Validity:

The credibility of a research design requires that the highest standards of validity and reliability be aimed for, even though each design, to some degree, often exchanges one standard for another. Validity requires that the identified variables in a research project be variables which exist in actual human events, and two types of validity have to be controlled for -- internal validity requiring that the data come from an authentic setting and external validity requiring that the findings be generalizable to other groups in similar settings.

In the present study, the primary control for internal validity is location of the investigation in actual inner-city schools and classrooms, and the selection of a wide variety of informants, including administrators, students, the teachers who were observed, and the observers who, for the most part, had teaching experience and who came from research backgrounds in anthropology, linguistics, English, and education.

The primary control for externality validity is the sampling procedure. First, the five schools selected for study were identified as inner-city schools in a number of government documents using typical indicators such as youth unemployment, number of students in EDY programs (Education of Disadvantaged Youth), and concentration of ethnic and racial minorities. In each of the five schools, minority student populations were 90% or more of the total school population, and in four of the five schools 15% or more of the students came from homes in which another language was spoken. In one school, 40% of the students came from homes with a second language.

Six criteria were used to select the fourteen local classrooms
with fourteen different teachers: (1) the classrooms were enrolling students from the inner city of the district, not some other area within the district; (2) they contained some writing instruction; (3) they were at various grade levels—9, 10, 11, and 12; (4) they had courses focusing on writing for students at different levels of writing competency -- remedial and bilingual versus general and advanced; (5) they had courses representing different ways of organizing instruction -- composition electives and journalism versus regular English classes; and (6) they had teachers with a reported diversity of approach and experience. The fourteen classes included two classes for advanced students, three for remedial, and two for bilingual. The other classes were four general English classes, two composition electives, and one journalism.

A second control for external validity was a sample of national teachers against which local teachers could be compared. In each of seven major cities in different parts of the country (Denver, New York City, Phoenix, Boston, Chicago, Los Angeles, and Memphis), twenty inner-city teachers ranked a set of anchor papers from a school district assessment and completed a two-part questionnaire on teaching experience and practices. One hundred and seventeen teachers completed the ranking and the questionnaire, providing information showing that local and national teachers were comparable in such matters as teaching hours, types of courses, range of student writing (the teachers estimated the number of students writing at a given level) and student attendance. However, local teachers had slightly less teaching experience than the national sample.

Reliability:

Reliability refers to the likelihood of obtaining the same answer
if one were to measure the same thing twice, and this design problem is the greatest challenge to the ethnographic researcher. Two measures are involved. The first measure involves internal reliability in the collection of data from various sources -- Do observers agree about what was observed or said? The second measure involves external reliability in the arrangement of the data into categories or constructs for interpretation -- Do observers agree about what the data means?

Data was collected from five sources -- interviews, writing samples, questionnaires, school documents, and over 300 full-period observations of the fourteen classrooms. The actual items in the questionnaires and documents were used for analysis, and, therefore, these sources presented no difficulties of internal reliability. However, the other three sources did present such problems. For example, because the writing samples had to be coded for type of student error (subject-verb agreement, pronoun case), the coder-observers prepared a list of errors and decision rules for each error, and practiced coding until agreement reached .9 or better. The reliability of interview coding was handled in a similar fashion, except that the data being coded in the interviews was much more open to differing interpretations and required fewer coders with more practice refining categories.

The internal reliability of the data from the classroom observations was checked and refined in two stages of practicing data collection. In the first stage, the fourteen observers and other members of the research staff practiced observations and coding of video tapes of classrooms, to develop an agreed upon set of terms and
format for observations notes and classroom layout, a coding matrix for classroom behavior and events during a given period (lesson structuring by the teacher, administrative work by teacher and students, disciplinary action by the teacher, and response interactions), a form itemizing traits of each writing episode during a class period, and a class summary partitioning the class period into its various segments (opening of class, lesson activity 1, lesson activity 2, and so forth). Decision rules for the various forms and terms were established in three weeks, and during the next three weeks the research staff practiced using the established routines, again on video tapes. A check of the reliability of the matrix coding of field notes showed an agreement on coding in 90% or more of the items.

In the second stage of practicing data collection, the observers field tested the procedures in two sets of classroom visits in different cities. In one visit, researchers worked in pairs so that the agreement in their data collection procedures and results could be compared and any discrepancies corrected. In summary, the classroom observations produced the following documents:

<table>
<thead>
<tr>
<th>Matrix of Behaviors</th>
<th>Segments/Lesson Summaries</th>
<th>Writing Episode Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field Notes + Class Layout</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

These documents became the basis for data analysis, which involves issues of external reliability. For example, the matrix provided a notation for each discipline event during a class period, and the total number of discipline events was calculated from these notations. In addition, because the matrix codes were marked in the left margin of the field notes, at the point where the behavior occurred, the actual
descriptions of the discipline events, including student behavior and teacher response, could be easily extracted from the field notes and grouped together for analysis. In the analysis of discipline, writing assignments, interruptions, teacher moves to structure a lesson, and other matters, copies of field notes were cut-up and arranged as groups of different kinds of behavior and events.

The first task in data analysis was to separate out the data on different variables in the field notes and in the coding documents (matrix, segment summary, writing episode summary) and group these variables in three levels of classroom context: (1) the writing assignment, including type of assignments given, amount of time spent writing, and variation of assignments by type of class; (2) the lesson, including teacher moves to segment the class period, opening and closing different activities in a lesson; and (3) the class period, including discipline episodes (student behavior and teacher response), handling interruptions, monitoring attendance, and coordinating resources.

The next step in data analysis was to categorize the discipline events into different teaching styles, the writing assignments into different types, the opening and closing of segments into different teaching styles, and so forth. Because the actual notes on an event were used for developing an analytical category, differences in analysis could be checked against the actual data.

A method of triangulation was used to help establish the external reliability of the data analysis. That is, different members of the research staff and different participants at the school level were asked to give their perspective on how to categorize and interpret the data. For example, the fourteen classroom teachers were asked to sort
the teaching instructions of teachers (from the field notes) into categories for assignment types. In addition to the teachers, both local and national, the other informants in the study were 117 students (same total as the national teachers) and twenty-nine other educators, primarily school administrators. The other educators were especially helpful in describing a fourth level of context -- school policies, including both institutional (testing, curriculum) and professional (learning, teaching) guidelines. The students gave their views of discipline, writing assignments, and a number other variables, including a student ranking of the papers scored by local and national teachers.

The study of policy issues influencing writing instruction was based on interviews of six local teachers and twenty-nine other educators assigned to nine different roles and four different administrative levels (state, county, school district, and school site). The twenty-nine other educators were named by various local teachers as sources of policies which influenced the teaching of writing.

These four different levels of context -- writing episode, lesson, class period, and policies -- were recognized by most informants as different roles played by teachers: (1) the teacher as a writing teacher; (2) the teacher as a teacher of language use in general, including reading and writing, giving and collecting the assignments, planning the lessons, responding to student writing, and establishing routines; (3) the teacher as a Manager, managing student attendance patterns, class interruptions, discipline problems, and resources for teaching; and (4) the teacher as a Policy Adviser/Implementer.
attempting to link policy and practice.

Results

The Writing Episode in all classes averaged 16 minutes and consumed one-third of the total class time. 71% of the 10th graders and 39% of the 12th graders wrote samples judged to be below the district's minimum standard for graduation from high school. The reasons for low scores tended to vary from one competency level to another, the papers at the bottom needing fluency, the papers in the middle needing appropriate audience focus, and the papers near the top needing more attention to organizational form. Only five of the eighteen errors rated by teachers as important showed a significant interaction with the scores on writing samples (ANOVA F from 3.39 to 12.94, p from <.000 to <.039).

Five kinds of assignments were found to occur in writing classes: (1) writing assignments based on reading ("Write about The Miracle Worker"), (2) assignments based on ideas ("Write as many memories as you can"/"Take notes on this"/"Write on one of these topics"), (3) skill assignments, varying from copying to exercises, (4) rhetorical assignments, providing an audience emphasis or an explicit form, and (5) a multiple focus, emphasizing more than one of the other four.

In all fourteen classes, skill assignments were the most frequent, and idea assignments consumed the largest proportion of total class time. However, teachers appeared to vary assignments to fit the needs of particular students. Remedial and bilingual classes had the largest proportion of class time devoted to idea assignments which helped with the students with fluency, their greatest need. Composition electives, on the other hand, had the largest proportion of class time devoted to multiple focus assignments, General English
classes had the greatest amount of assignment diversity, and Advanced English classes had the smallest proportion of total class time devoted to writing. Students in Advanced classes allegedly needed less on-task assistance from the teacher.

The adaptations for various classes were also evident in the way the teachers segmented their Lessons as they shifted from one activity structure to another. The teachers almost always segmented their class period into more than one lesson activity, the average for the fourteen classes being 2.1%. However, teachers of bilingual students and ninth graders segmented more than other teachers, averaging 2.9 (the highest being 3.4). The reason given for this lesson strategy was that these students had the greatest tendency to drift off-task and activity shifts were ways to bring these students on-task. The segments in twelfth grade classes averaged 1.9. The classes with few segments tended to occur with older groups, particularly 12th graders, both in Advanced and Remedial classes, suggesting that older students or those who do not drop out can stay on-task longer.

Despite the evidence of teachers varying assignments to meet the needs of different students, skill assignments were persistent and frequent in all classes. Short answer assignments (focusing on letter, word, phrase, sentence) were the most frequent assignments (41%), the shortest (10 minutes), and the second largest percentage of total class time (27.3%). This is comparable to the trends reported by Applebee (1981) and lamented by Cooper (Preface to Applebee).

Why this emphasis on skill assignments? Two sets of conditions appeared to give a high degree of "survival value" to skill assignments. The first condition was the flow of management problems
during a class period, a flow which will back-up and later become a potential flood if the teacher does not attend to these problems. These problems took the form of 3.19 discipline problems each class period, 1.71 interruptions (school bulletins; messages over the telephone and from student messengers, usually from counselors or the principal’s office; events in the halls; the public address system), attendance (taking roll with an average of 30% absent in each class and monitoring/recording .74 late students each class period), and coordinating classroom resources (paper, pencils, other equipment).

Skill assignments, because they required a minimum of beginning directions and supporting instruction, gave the teacher the necessary time to set aside instructional responsibilities and turn to management problems. Skill assignments were also an efficient strategy for handling a number of policy conditions which were part of the district’s testing and categorical aid programs. Skill assignments focused on small, easily measurable units of instruction which often paralleled the objectives listed in the district policy documents on curriculum and accountability. Reporting the number right and wrong on an exercise was much easier than reporting development in fluency, audience awareness, or organizational skills.

These are only a few of the results of the study of the teaching of writing in inner-city secondary schools (Myers and Thomas, 1982). But these results suggest that good instruction in writing requires an understanding of more variables than those found in decontextualized studies of writing. These studies are useful, but they do not give teachers lessons for classrooms. Lesson design is more complicated than that. What is needed are more studies of great teachers handling the many variables in their classrooms, and ethnography is a promising
methodology for meeting this need.

*In a large study such as this, a number of people make critical contributions. Four need to be named here: Susan Thomas, who acted as coordinator of the research assistants, and the three members of the Advisory Committee -- Robert Ruddell and Walter Loban from the University of California, Berkeley, and Arthur Applebee from Stanford University.*


Reed, Rodney. "Urban secondary school leadership: Contemporary


