DOCUMENT RESUME

ED 397 126 TM 025 269

AUTHOR Flores, Kathryn Younger
TITLE Measuring Instructional Effectiveness: A Comparison of a Computer-Assisted Systematic Observation Instrument with Global Measures. Proceedings from Seminar on Teacher Development and Linguistic Diversity.
INSTITUTION Southwest Regional Lab., Los Alamitos, CA.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE Oct 95
CONTRACT RR91002006
NOTE 48p.
PUB TYPE Reports - Research/Technical (143) -- Collected Works - Conference Proceedings (021)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Classroom Observation Techniques; College Faculty; Comparative Analysis; *Computer Assisted Testing; Computer Software; Evaluation Methods; *Evaluators; Higher Education; *Instructional Effectiveness; Measurement Techniques; Music; *Performance Based Assessment
IDENTIFIERS *SCRIBE Computer Program; University of Texas Austin

ABSTRACT

This paper presents preliminary, but statistically significant, findings from a study that compares two methods of measuring instructional effectiveness: global evaluation by experts and systematic observation using the SCRIBE Ob.2 software developed at the University of Texas at Austin. Hierarchical instruction of a performance skill (instrumental music) was the topic of study. Eight evaluators, four from music education and four general education evaluators, ranked six videotaped rehearsal segments in order of effectiveness and provided brief comments regarding the criteria used to rank the segments. Independently the same excerpts were analyzed using the SCRIBE software. The SCRIBE observation instrument was used to record data, and the software allowed for configuration of subjects and behavior categories. Judges from within and outside the field of music education consistently identified exemplary and substandard teachers, but rankings of teachers of intermediate effectiveness were less reliable. The SCRIBE software was a useful and generally manageable means for gathering information on student-teacher interactions. It identified ways in which exemplary and less effective teachers differed to a statistically significant degree. This pilot study holds promise for improved evaluation approaches. (Contains 4 tables, 13 figures, and 23 references.) (SLD)

Southwest Regional Laboratory

4665 Lampson Ave., Los Alamitos, CA 90720

(310) 598-7661

Proceedings From Seminar on Teacher Development and Linguistic Diversity

Kathryn Younger Flores

Southwest Regional Laboratory

October 1995

Proceedings From Seminar on Teacher Development and Linguistic Diversity

Measuring Instructional Effectiveness: A Comparison of a Computer-Assisted Systematic Observation Instrument With Global Measures

Kathryn Younger Flores

Southwest Regional Laboratory

Deliverable 21.3-94

October 1995

Prepared under a subcontract with Far West Regional Laboratory for Educational Research and Development (contract no. 91002006, U. S. Department of Education). The content does not necessarily reflect the views or policies of the Far West Laboratory or the U. S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by these agencies.

The Southwest Regional Laboratory (SWRL) is a nonprofit, public educational agency that exists to address challenges resulting from changing demographics and increasing numbers of children placed at risk in the Metropolitan Pacific Southwest. The Laboratory addresses its mission by engaging in research, development, evaluation, training, technical assistance, and policy analysis.

Contents

Tables

Figures

Abstract

This paper presents preliminary, but statistically significant findings from a study that compares two methods of measuring instructional effectiveness: (a) global evaluation by experts and (b) systematic observation using the SCRIBE software developed at the University of Texas at Austin. Hierarchical instruction of a performance skill (instrumental music) is the topic of the study. Instrumental music was chosen for the for two reasons: (a) the highly observable and easily quantifiable nature of the student and teacher behaviors and (b) the high degree of success reported by instrumental music programs for the mainstreaming of diverse and differently able learners.

Eight evaluators (four from within the field of music education and four general education evaluators) ranked six videotaped rehearsal segments in order of effectiveness and provided brief comments regarding the criteria used to rank the segments. Independently, the same excerpts were analyzed using the SCRIBE software. Independent variables coded for the systematic observation were selected based upon previous studies using similar procedures. Statistical procedures and graphic comparisons were used to compare the rehearsal excerpts by rank.

Judges from within and from outside the field of music education consistently identified exemplary and substandard teachers, but ranking teachers of intermediate levels of effectiveness were less reliable. The SCRIBE software was a useful and generally manageable means for gathering information on student-teacher interactions. Exemplary teachers differed to a statistically significant degree from less effective teachers in their frequency of specific task directions, frequency of specific approvals and disapprovals, greater percentages of time spent modeling desired musical outcomes, time spent in question and answer exchanges, and in the sequencing of instructions and feedback. The sequencing of feedback had a significant effect upon rank.

Although no extraordinary modifications were used to accommodate diverse learners, experts posed some new avenues for exploration: (1) the use of written advance organizers for the rehearsal was correlated with higher ranking, (2) experts called for more feedback from peers, and (3) more use of peer observation and assistance. The high reliability of expert global evaluation combined with the greater objectivity and specificity of systematic observations can be used to improve the use of evaluations for prescriptive purposes and to lend greater validity to the independent variables used in empirical studies of instructional effectiveness for diverse learners.

Introduction

This paper presents preliminary, but statistically significant, findings from a study that compares two methods of measuring instructional effectiveness: (a) global evaluations by experts and (b) systematic observation using the new SCRIBE Ob. 2 software. Hierarchical instruction of a performance skill (instrumental music) is the topic of the study. This study demonstrates that the systematic observation method is a useful and generally manageable means for gathering information on student-teacher interactions. Although no extraordinary modifications were used to accommodate diverse learners, commentaries from both music educators and general education evaluators posed some new avenues for exploration, which are discussed in the conclusion of this paper.

Identifying Variables Associated With Effective Instruction

Identification of the variables associated with effective instruction, and application of this information to classroom practice, has been of long-standing interest to the educational research community. Rather than merely discriminating between good and poor teaching, evaluations should offer teachers target objectives for improvement, as well as workable strategies for attaining the desired outcomes. Evaluations cannot be based on isolated examinations of teacher attributes or student performance. To pinpoint elements of effective instruction, observations should be made of the interactions between students and teachers during lessons (Flanders, 1964, 1970; Sang, 1982).

In spite of the potential usefulness of observational data, relatively little research has been published using systematic observations of student-teacher interactions because agreement has not been reached regarding the validity, reliability, and practicality of the methods used for conducting and reporting this type of study (Shaver, 1983). In addition to the methodological concerns, systematic observation has been a very cumbersome procedure. Recent advances in computer technology have greatly simplified the data collection and analysis procedures. Researchers now are better equipped to attack the problem of identifying valid and reliable indicators of effective teaching.

Global ratings of instructional effectiveness are used widely for evaluation. Although often reliable and convenient (National Evaluation Systems, 1994-95; Siebenaler, 1992), rankings and ratings usually are not helpful for prescriptive purposes. This is because global ratings by expert evaluators generally do not distinguish among the program, personality, and implementation variables contributing to the teacher's score or rank order. Furthermore, without specific behavioral information, teachers may find rank and rating numbers somewhat useless for planning improvements (if not downright disheartening).

Systematic observation can be used to obtain specific data (Duke & Madsen, 1991; Siebenaler, 1992). Although not always the method of choice, systematic observation is particularly useful for observing hierarchical instruction of observable performance or motor skills (Fitts & Posner, 1967; Gagne, Briggs, & Wager, 1988). Comparisons of systematic observation data to global evaluations and commentaries by expert judges may contribute to understanding the variables and corresponding measurements associated with effective instruction so that teacher assessment information can be used for prescriptive as well as evaluative purposes; and to lay a more secure foundation for empirical research by providing reliable measures of target behaviors.

Perspectives

Although experts can differentiate between good and inadequate instruction with a high degree of reliability, identification of the specifics of effective teaching remains elusive. The numerous variables and the interactions among teacher, student, and subject matter variables complicate the study of teaching and learning. Not surprisingly, agreement has not been reached on an operational definition of effective instruction for the purposes of prescriptive evaluation and experimental study.

Attempts to explain the complexities of instruction by recording a handful of selected behaviors have been criticized because reductionistic studies negate the "art" of teaching. It is not difficult to see why evaluations have traditionally consisted of experts' general perceptions of effectiveness. Global evaluations view instruction as a synergistic entity—as more than the sum of the parts. In spite of the relatively high reliability of expert opinions and convenience of the method, global evaluations fail to spell out ways for the teacher to maintain or improve instructional quality. In addition to the lack of specific information provided to teachers, anecdotal reportings and global ratings can lack sufficient validity for empirical study of program effectiveness.

On the other hand, behavioral checklists and systematic observation instruments focus the observer on a predetermined set of criteria that may or may not reflect the truly important aspects of teaching (Brophy & Good, 1986; Duke & Prickett, 1992). Another disadvantage of reliance upon a set of "effectiveness indicators" is that assessments of teaching quality tend to be idiosyncratic or at best, culturally determined. Due to rapidly changing classroom demographics, evaluators are less likely to know what to look for in unfamiliar situations (Wright, 1995). The experiences of students may be vastly different from those of the evaluator. Furthermore, expectations for how instruction is to be conducted may differ radically from one culture to the next.

Movements to educate as many students as possible in regular education classes call for changes in instructional strategies. Teachers may not receive adequate preparation and or assistance to accommodate students with disabilities and or those students with limited exposure to the English language and dominant school culture (Obiakor, 1992). Teaching as one has been taught or as one has been teaching over the years may do little to avert the drop in standardized test scores, the increasing demands of classroom management, and the problems that affect students' motivation to learn. Though poor student performance scores may not entirely reflect the quality of instruction, lack of student achievement can have a negative effect upon teacher and program evaluations. Because many of the variables contributing to student achievement are

beyond the teacher's control, teachers must actively seek ways to maximize effective use of instructional time.

The study of instrumental music instruction is convenient because of the highly observable nature of the behaviors and is particularly interesting because of the often paradoxical results of music education. In light of the pleasurable nature of music and the initial enthusiasm of many beginning musicians, it is somewhat surprising that so few adults continue to pursue music performance for enjoyment. In spite of the highly reinforcing nature of music, many novice performers abandon their studies of a musical instrument before they ever reach the necessary level of independence (Lawrence, 1978).

The teacher's responsibility goes beyond motivating students during rehearsals. Students must know how to practice effectively at home. Fewer students, regardless of ability level, receive private lessons or instruction at home on their musical instrument. Changing demographics present challenges to teachers who can no longer afford to choose only the most able or those who can afford private study. Teaching as one has been taught, without regard to monitoring program effectiveness, is unlikely to produce unprecedented numbers of enthusiastic musicians from an increasingly diverse student population. Strategies must be found to prevent the frustration and hopelessness that lead to attrition from music programs, if participation in musical ensembles is to become a realistic goal for more students.

Theoretically, any subject matter or body of knowledge should be accessible to virtually all learners if presented in an orderly manner (Gagne, 1977). Playing a musical instrument is one example where a complex sequence of motor and intellectual tasks must be mastered in order to fully appreciate the activity. Each skill must be learned and repeated many times in a variety of contexts. Mastery of an instrument requires patience, but excessive repetition of technical drills without playing real music or without improvement in the performance quality, or both, becomes extremely punishing to the learner (not to mention the teacher and audience). Unfortunately, striking a proper balance between technical exercise and uninterrupted playing time remains an elusive competency to many teachers.

Presenting a learning hierarchy such that tasks of increasing difficulty are mastered over time contributes to the student's sense of self-efficacy and lead to improved performance of skills (Gagne, 1977). In fact, longitudinal studies of highly accomplished individuals in the fields of music, sports, mathematics, and neurology indicate the great importance of technical guidance and perceived progress (Bloom, ed. 1985). Furthermore, by following the steps modeled by the classroom teacher, the student may replicate good learning strategies during independent practice. For this reason, learning hierarchies should be structured and implemented such that students experience more successful and motivating performance episodes. In addition to sequencing instruction in the curriculum, to be effective, teachers must offer the proper guidance

in terms of clear directives, modeling desired behaviors, making accurate discriminations, and offering useful feedback (Sang, 1982) during rehearsal time. Systematic observation may provide a way for teachers to examine and improve how they present instruction and offer feedback to students. Managing the directive and feedback variables would help teachers optimize instructional time, while leaving sufficient time for uninterrupted playing.

Purpose of Study

The present study seeks to operationalize Gagne's theories (Gagne, 1977; Gagne, Briggs, & Wager, 1988) on sequencing in instructional design, by using a temporally sensitive, computer-assisted, systematic observation instrument (SCRIBE), developed at the University of Texas at Austin, School of Music. Information from the research will be used to:

- observe the frequencies, rates, durations, and sequences of selected student and teacher behaviors as they occur in time; and
- explore the nature and extent of the relationship between systematic observation data and independent, global evaluations by experts in the field.

The purpose of this study has been to observe systematically student-teacher interactions as they occur in time during school band and orchestra rehearsals. Plans have been made to refine the methodology for future use in more detailed observational and possibly empirical research. The present research in no way implies that the higher order aesthetics of music can be reduced to exact measurements of teacher and student behaviors. Nor does it imply any similar applications or limitations to the art of teaching. It is known, however, that examining the variables found to have an effect upon learning, transfer, and retention may aid the sequencing of instructional events to allow for a more positive learning experience. With this information, teachers can be more proactive to ensure positive learning experiences for all students (Duke & Madsen, 1991).

Previous studies in music education (Duke & Madsen, 1991; Price, 1983; Sang, 1982; and Siebenaler, 1992) have measured the frequencies, durations, and rates of:

- teacher directives (instructions, modeling, cues);
- student performance (correct or incorrect and quality);
- teacher feedback (approvals, disapprovals, specific and nonspecific); and
- task direction (whether successive directives move ahead, backtrack to teach subskills, or repeat the same step).

Summaries of these measures have been examined in relation to teacher effectiveness and student attitudes. In addition, Siebenaler (1992) has examined the sequencing of teacher and student behaviors during private piano instruction; and Bolte (1994) has observed the sequencing of events during chemistry instruction, using the Kieler Observation Instrument. The literature documenting the temporal aspects of music rehearsal interactions is scant. The observation procedure for the present study describes teacher and student behaviors in terms of the frequencies, durations, rates, and sequencing of teacher directives, feedback, and task direction in relation to student performance as the behaviors occur in time. Data from systematic observation are then compared to independent, global measures of teacher effectiveness as determined by a panel of experts in an attempt to isolate any variables found to be related closely to measures of effective instruction.

Research questions. The present study seeks to determine:

1. Is the SCRIBE systematic observation instrument a manageable tool for use with large and small group musical instruction, and how can it be used most effectively for large group observations?

2. What is the nature and the extent of the relationship between global evaluations of teacher effectiveness and systematic observation measures of teacher and student behaviors and lesson progress?

3. In light of the current emphasis placed on multicultural and language development methodologies for all teachers in California, will expert observers cite new or different categories of behaviors to identify exemplary instructional practices? If so, will the teacher and student behaviors indicative of effective instruction differ significantly from those currently documented in the existing body of literature on music education research, thus requiring modifications in the selection of observation categories?

Method

Type of Study

This was an observational pilot study. No differential treatments served as independent variables among groups. All data analyses were exploratory.

Participants

The 6 participants were chosen from a larger original data set of 12 instrumental music teachers in order to depict as much contrast as possible. The teachers were filmed during

rehearsals with students from a wide range of socioeconomic levels and ethnolinguistic groupings. Public school directors of intermediate level, secondary school instrumental ensembles in Los Angeles County were invited to serve as research subjects for the present study. Twelve teachers were videotaped during three consecutive rehearsals of the same ensemble. One additional participant was a professor of music education and renowned ensemble director who was filmed during a guest appearance with an intermediate level ensemble in Austin, TX. Except for the guest conductor, all teachers were full-time employees of their respective school districts and were certified by the state of California to teach music. Nearly all had completed some type of advanced studies beyond the master's level, including state mandated training for teachers serving students with exceptionalities and those designated as limited English proficient (LEP). Each participant had over five years of experience as an ensemble director. Five males and one female were presented on the sample tape. Their ages ranged from 28-58 at the time of the study. Music directors included in the pilot study were of African American (2) and White (4) ethnicities. Participation was voluntary and no attempts were made to control demographic factors. Willingness to participate in the research study may have self-selected those directors who felt more confident about their teaching or musical ability, or both.

Each teacher was asked to select his or her intermediate level ensemble. Three consecutive rehearsals were videotaped to minimize the effects of chance occurrences and the introduction of an observer with a video camera into the environment. Teachers were told that the purpose of the study was to observe the normal, everyday interactions taking place during the course of rehearsing a piece in progress. The music directors were simply told to proceed as usual.

Setting

All lessons were recorded in the usual classroom or rehearsal hall setting. The ensembles represented the variety of instrumental music classes at the intermediate level that could be found throughout the greater Los Angeles area. Permission to conduct the study was granted at the district or school site level, as required. The students ranged from 11-16 years of age, and represented middle (grades 6-8), junior (grades 7-9), and senior high (grades 9-10) school students. All socioeconomic levels were represented in the sampling. Ensembles reflected demographic profiles typical of schools in Los Angeles County, including mainstreaming of special education students and ethnolinguistic diversity across all socioeconomic levels. Class sizes ranged from 20-50 students.

Intermediate level ensembles were selected to represent the most strategic opportunity to view interactions directed at improving student performance through intensive technical

assistance. Differences in experience were evident. Students who had the opportunity to begin studying an instrument at an early age clearly were more advanced than those who began playing as secondary students. In some cases, access to private instruction may have accounted for differences in musical level. Teachers reported years of study to range from 1-12, but were unable to specify length of study for individual students. Even within groups, wide variations were present. Placements into intermediate ensembles often were made on the basis of scheduling convenience and the lack of other available electives, rather than on present level of musical achievement.

Conditions of space, room arrangement, lighting, noise, and interruptions varied among the settings. Several locations were less than ideal for filming in terms of space, lighting, and noise. The availability of only one camera also limited visibility at any given moment. The environments were left untouched in an attempt to represent the situations as they would occur had the observer not been present. Lessons were not manipulated in any way.

All equipment was set up prior to instruction to avoid unnecessary commotion. The observer was unable to leave the video equipment unattended, but movement about the room was kept to a minimum to avoid distracting the students and teacher. The observer made every attempt to remain as unobtrusive as possible.

Observational Procedures

Videotaping of the individual lessons for the study was done between May 1994 and February 1995, using a Panasonic 170 VHS camera on a tripod. No additional lighting or sound amplification was used. The rehearsals or classes lasted approximately 55 minutes. The participants were told to proceed as usual. The observer explained that she would not interact in any way with the students, but would answer questions or concerns. Students were assured that videotapes were to be reviewed exclusively by researchers, but were informed that they had the right to decline to be filmed. All students agreed to participate.

For analysis, the videotaped segments initially were screened to find the portion of the lesson that would best illustrate the behaviors under study. Wherever possible, a piece or exercise in progress was viewed over the three-day period. The lesson excerpts had a mean duration of 10 minutes, ranging from 8-12 minutes. Not all observations and analyses of the performance segments extended to the conclusion of the piece. Videotaped portions were edited such that the selected interactions provided an adequate sampling for analysis. Each observation began precisely with the teacher's initial instruction to the group as they were preparing to attempt the particular musical task. In the event of an unclear beginning, the most logical

starting point was chosen for establishing a consistent reference for repeated observations. A logical point of closure was inserted to end each selected excerpt.

After viewing the tapes, no extreme contrasts between effective and less effective teachers were apparent to the expert group. In the interest of expediency, a tape was included of an internationally acclaimed music educator rehearsing a middle school band in Austin, TX, for the following reasons:

- to provide higher contrast of effective and less effective direction;
- because far less is known of the behaviors associated with effective teaching than with ineffective teaching (Siebenaler, 1992);
- global evaluations of good and acceptable music instruction have been far less reliable than those of inadequate instruction on inter-rater measures (Siebenaler, 1992), though a study of macro-measures of academic instruction revealed greater consistency in the identification of good teaching (Swank et al., 1989); and
- because evaluations of music educators often are carried out by nonmusicians or by persons unfamiliar with music education, it would be particularly helpful to identify some behaviors associated with effective rehearsal technique that would be apparent to experts outside of the music education field.

A new sample tape of six excerpts representing contrasting instructional presentations was edited for analysis. Segments were chosen to depict contrast in teacher activity, intensity, and strategies. The sample tape of six excerpts was viewed by a small committee of eight expert pedagogues from the Southwest Regional Laboratory; California State University, Long Beach; and the Los Angeles Unified School District. Each evaluator was instructed to view the tape, rank order the samples from most to least effective, and comment on what they found particularly effective and ineffective in the excerpts. Evaluators were free to view the film as often as they wished. Concurrently, data were collected by the principal investigator using computer-assisted systematic observation procedures for each edited lesson segment.

Data Recording Procedures

Recent advances in computer technology have greatly simplified systematic observation procedures. The SCRIBE observation instrument was used to record data on a Macintosh Classic or SE/30 computer while viewing the videotapes on standard, household video equipment. The C++, based SCRIBE software allowed for configuration of subjects (participants) and behavior categories. The observer used a point-and-click interface to record selected behaviors in each category. Data were recorded during multiple passes of the videotape.

Durations of behaviors were recorded for the length of time the mouse button was depressed. At the conclusion of the observation interval, the program displayed the frequency, total time, percentage of total time, mean, standard deviation, and rate of the selected behaviors, and reported the timed sequence of events. Selected behavior categories were reviewed by another trained observer to verify reliability.

Development of Observation Categories

The behavioral categories used in this study were chosen on the basis of the existing body of knowledge on classroom observations. In an attempt to select the categories most likely to yield answers to the proposed questions, research on music education and systematic observation was reviewed with particular attention to the isolation of variables relevant to large-group, skill-based instruction. Because of the nature of the populations used for this study, additional categories were chosen based upon recommended practices for secondary, content area instruction of LEP students (Richard-Amato, 1988; Tharp & Gallimore, 1988).

Existing systematic observation forms were reviewed for applicability to the present study. Of particular interest were instruments that included measures of lesson progress (Duke & Madsen, 1991) and temporal measures of the instructional sequence (Bolte, 1994; Siebenaler, 1992). By comparison, the categories developed by Bolte (1994) for computer-aided observation of chemistry instruction were considered for relevance to skill-based instruction. With the exception of the study of chemistry instruction, the bulk of the extant research involved individualized, rather than large-group, instruction.

Research in language development strategies for secondary school students in instrumental music classes is scant. The California State Curriculum Frameworks recommend offering second language learners ample opportunity to practice language skills in a natural context. Because of the scarcity of research in bilingual, instrumental music instruction, whether increased discussion between students and teachers has any significant effect upon musical performance, student attitude, motivation, or English language proficiency remains unclear. However, Siebenaler (1992) and Bolte (1994) found questioning of students to vary inversely with student attitude. Music learning has been the primary focus of the present study, but language development behaviors were considered in light of commentaries from the expert judges.

To organize the data, observation categories have been grouped according to Teacher Behaviors, Student Behaviors, and Lesson Progress as follows:

Teacher behaviors. By surveying the extant literature, the teacher responses deemed significant and observable, and have been recorded as follows:

1. Task Instruction (T): Specific information on the manner in which a task must be performed, including technical and expressive recommendations: "Play softer at B, raise your second finger on the F#."

2. General Instruction (G): Information of a nonspecific or organizational nature: "Try again, start at B, etc."

3. Specific Assistance (SA): Guidance or assistance other than verbal, such as raising an elbow, moving a finger, opening a book to a page, tuning an instrument, or pointing to notes for the student.

4. Music Talk (MT): The teacher explains some aspect of the music, including lecture or response to questions on all facets of music not specifically related to how a task step is to be executed.

5. Sing/Clap (S/C): The teacher models the desired behavior or musical effect by singing, clapping, or tapping. Normal conducting gestures are excluded from this category.

6. Modeling (M): The teacher models the desired musical effect by performing the example on a musical instrument.

7. Play (P/WS): The teacher accompanies the ensemble or performs with the ensemble.

8. Off-Task (OT): The teacher behavior is unrelated to the learning task or to the music.

Teacher feedback has been coded according to the following categories:

1. Specific Approval (A);

2. General Approval (a), for example, "Good, OK";

3. Approval Mistake (AM): The approval is either erroneous, or it is not contingent upon the preceding student response;

4. Specific Disapproval (D);

5. Nonspecific disapproval (d): Negative, general commentary; and

6. Disapproval Mistake (DM): The disapproval is inappropriate or unrelated to the preceding student behavior.

The preceding categories were coded to represent salient features of the instructional portion of the student-teacher interactions. Teacher talk in a language other than English was not coded because no substantive instruction took place in another language.

Student behaviors. By surveying the extant literature, the student responses deemed significant and observable, and have been recorded as follows:

1. Play (P): Entire group plays.
2. Smaller Group Plays (SG).
3. Individual Plays (IP).
4. Sing/Clap (S/C): The student performs by clapping, tapping, or singing.
5. Questioning/Answering (Q/A): The student requests clarification or elaboration of any aspect of the learning task or the music, or answers a teacher's question.
6. Peer Assisting (PA): One student assists another student or the ensemble by explaining or translating a teacher directive or feedback comment, or by offering original feedback.

Student responses have been coded simply as (+) for correct and (-) for incorrect, contingent upon whether the majority of students performed the task requested by the teacher.

Lesson progress. Perceptions of success in mastering a learning task often are related to observable measures of progress. During the course of a rehearsal, progress may be measured by the rate at which the students move on to the next task step. Conversely, an excessive amount of backtracking and repetitions beyond those necessary to retain a skill may be viewed quite negatively by the students and teacher. The frequency and sequence of progress episodes can be used to clarify the correspondence between teacher behaviors and the students' ability to master a given musical task. By comparing the level of complexity of a task to the level of complexity of the preceding task, student progress has been grouped according to one of the following four categories:

1. Forward (→): The new task moves on to a new section of the piece or adds some level of complexity (increasing the tempo, putting sections of the ensemble together, going from tapping a passage to playing it on the violin).
2. Backward (←): The new task reteaches or simplifies the skill, by breaking it down into subskills, or decreases in complexity.

3. Repeat (/): The new task is the same as the preceding task.

4. New Task (X): The new task is unrelated to the previous task.

Independently, the six lesson excerpts were viewed by a panel of experts. Four of the judges were experienced evaluators in the field of music education (instrumental music specialists) and four were professional evaluators from an educational research and development laboratory. In addition, the evaluators met the criteria of experience or training, or both, in instructional methods for multicultural student populations. The overall quality was ranked for each lesson, and evaluators commented on the strengths and weaknesses of the segments receiving the highest and the lowest rankings. Systematic observation measures were compared by statistical procedures to evaluators' rankings and commentaries for the selected samples.

Findings

The data were analyzed in two parts. First, the interjudge agreement was determined for the rank ordering of subjects in the excerpts. The selection of excerpts was carried out to maximize contrast rather than to reflect the director's true effectiveness. Only the excerpts shown to the judges were observed systematically. Therefore, results do not necessarily reflect the rankings that might have been obtained by other observation or analysis procedures. The limitation was an intentional, cost-saving measure for two reasons: (a) little previous work had been done applying the new instrument to large-group instruction; therefore, recording the systematic observation data may have required significant modifications; and (b) interjudge agreement was crucial to further study.

The interjudge agreement was determined using the Kendall coefficient of concordance (W), a nonparametric procedure for analyzing ordinal level data. All analyses were run on SPSS for the Macintosh. Table 1 shows the results for overall ranking of subjects and agreements on the most and least effective excerpts.

Table 1
Results of Interjudge Agreement

Kendall Coefficient of Concordance for all six participants

Mean rank	Variable
3.13	SUB1
1.38	SUB2
2.50	SUB3
5.88	SUB4
3.50	SUB5
4.63	SUB6

Cases	*W*	x^2	df	Significance
8	.7179	28.7143	5	.0000

Kendall Coefficient of Concordance using the segments ranked as most and least effective

Mean rank	Variable
1.00	SUB2
2.00	SUB4

Cases	*W*	x^2	df	Significance
8	1,0000	8.0000	1	.0047

table continues

Kendall Coefficient of Concordance for the two most and two least effective segments

Mean rank	Variable
1.25	SUB2
1.88	SUB3
3.88	SUB4
3.00	SUB6

Cases	*W*	x^2	df	Significance
8	.8188	19.65	3	.0002

In addition to rank ordering film clips, the judges were asked to comment on the salient features that distinguished effective from ineffective instruction as depicted in the excerpts. Seven of the eight judges agreed upon the most effective sample. The one judge who was a professional evaluator with no musical training preferred the second ranking excerpt because of the subject's use of a written advance organizer, soliciting of student input for the evaluation of performance, and a democratic rather than authoritarian classroom climate. These qualities were mentioned as favorable by four (50%) of the eight raters.

All judges thought that it was relatively easy to spot the most and least effective segments, but that ranking the intermediate lesson sequences was virtually impossible. Qualities describing instruction are summarized in Table 2.

Table 2
Qualities of Effective and Ineffective Lesson Segments

Effective	Ineffective
Intensity; enthusiasm; excitement; personality	Disinterest; lacks "presence"; very dry; no charisma; just playing through the pieces
Concentration, focus on specific goals	Lacks structure, jumps from one activity to another for no apparent reason; long period of off-task time looking for papers

table continues

Effective	Ineffective
Clear or written objectives	Objectives either not stated or never followed up in practice; teacher is just kind of reacting to whatever happens and doesn't plan ahead (reactive rather than proactive)
Rapport with students; interacts with students; Maintains good eye contact; body language	Uninvolved with students; ignores students No eye contact except with score; tense or angry gestures and body language
Atmosphere reflects hard work, but is positive	Atmosphere lacks signs of motivation; attitudes range from bored to hostile
Students treated with respect; student input encouraged	Insulting students; focusing on irrelevant details (e.g., how to sit in a chair)
Good mix of approvals with constructive criticism	Comments usually are negative; lack of specific solutions to fix trouble spots
Feedback leads to audible improvements; excerpt ends on very positive note—"now you're playing like one instrument..."	Disapprovals get worse as time goes on, then teacher gives up and goes to something else, ignoring what really needs to be "cleaned up"; gives incorrect technical information
Feedback generally is positive	Feedback is negative or even insulting
Students play through entire piece, but time of playing segments is varied to allow for fixing details	Too many stops in one case and too few in the other
Easy to follow beat and gestures	Marking time, but not really conducting; can't tell what teacher wants by gestures; unclear beat pattern
Class is very attentive	Not out of control by any means, but significantly more disruptions and several students in the room not even playing
Ensemble plays in tune	Terrible intonation that doesn't get corrected

table continues

Effective	Ineffective
Attention to dynamics, articulation, and phrasing improve performance	No apparent improvement in dynamic contrast, articulation, or phrasing
Good musical choice	Poor choice of music

Both musical experts and educational evaluators made music and nonmusic related commentaries. The music experts generally were more concerned with musical improvements for the overall ranking, but three of the four nonmusicians reported similar criteria in less detail. Musicians and nonmusicians commented on personality factors, but only music educators commented on conducting gestures. The professional evaluators commented more on student behaviors or factors related to classroom climate.

Independently, the excerpts were analyzed using the SCRIBE observation instrument. Appendix A shows the actual output and profile for the highest ranked excerpt, and Appendix B, the lowest ranked excerpt. Reliability was checked for two tapes on five measures of student and teacher behaviors. The frequencies and durations of progress segments were rejected from the analyses because it was virtually impossible to determine what the teacher had in mind without the observer having a written script or even a lesson plan. Table 3 shows the reliability correlations for the selected categories of behaviors.

Table 3
Reliability of Observations by Two Observers

Behavior	Interobserver reliability
Student play frequency	1.00
Mean playing time	.85
SD of playing time	.93
Task directive frequency	.97
Music talk mean time	.82
Music talk *SD*	.90

For prioritizing the variables associated with effective instruction, the results of student and teacher behaviors and progress were correlated with subject rank. This procedure could not be taken as a measurement of the exact magnitude of the effect of each variable because of the

possibility of errors inherent in correlations of small samples with numerous and interacting variables, but was useful for preliminary comparisons. The mean time used for question and answers was negatively correlated with effectiveness. The rate of sing/clap episodes correlated with effective teaching. Correlations worthy of further investigation in order of priority are shown in Table 4.

Table 4
Comparison of Correlations for Selected Behavior Categories With Subject Rank

Behavior category	Correlations
Mean question/answer time	-.93
Rate of teacher sing/clap time	.89
Frequency of specific disapprovals	.80
Whole group play episode frequency	.80
Total question/answer time	.80
Frequency of task directions	.77
Small group total playing time	.71
Specific approval frequency	.70
Total percentage of task direction time	.66
All individual play categories	.66
All student sing/clap categories	.66
Mean off-task time	.66

Because of the lack of consistency between even the top two and bottom two subjects across all measures, graphic profiles are provided (see Figures 1-13), comparing subjects on selected variables.

Figure 1
Profile of an Effective Lesson Sequence: Teacher Directives

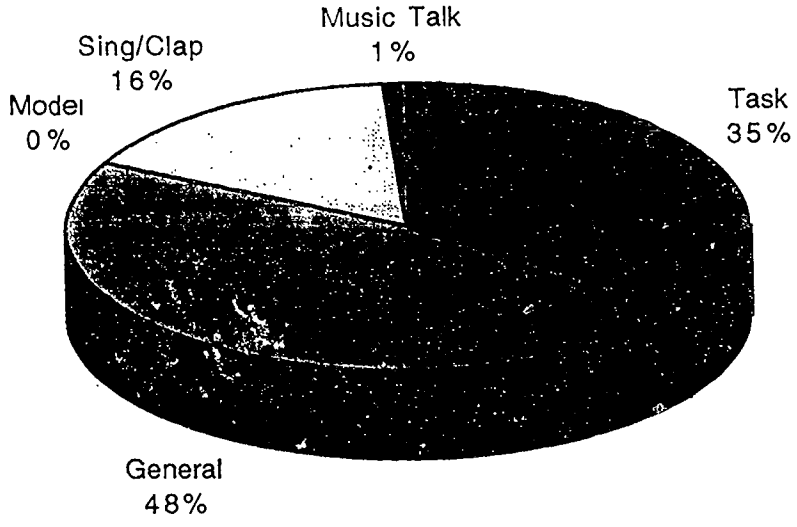


Figure 2
Ineffective Percentage Frequency of Directives

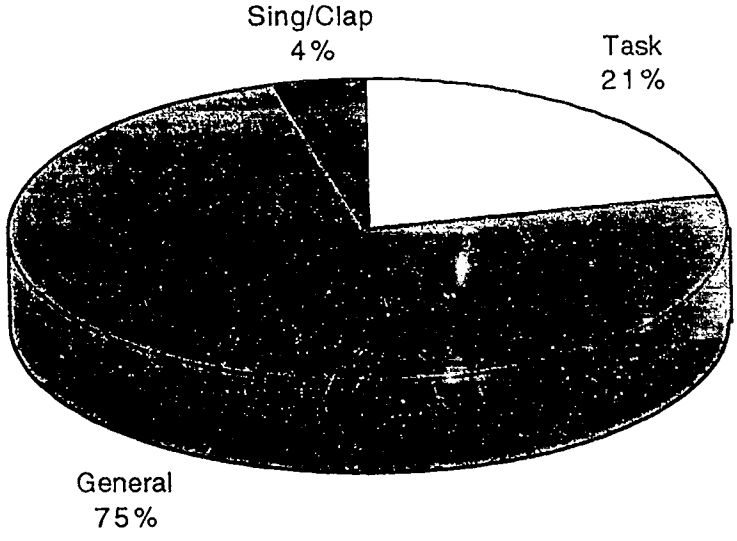


Figure 3
Effective Percentage Feedback Categories

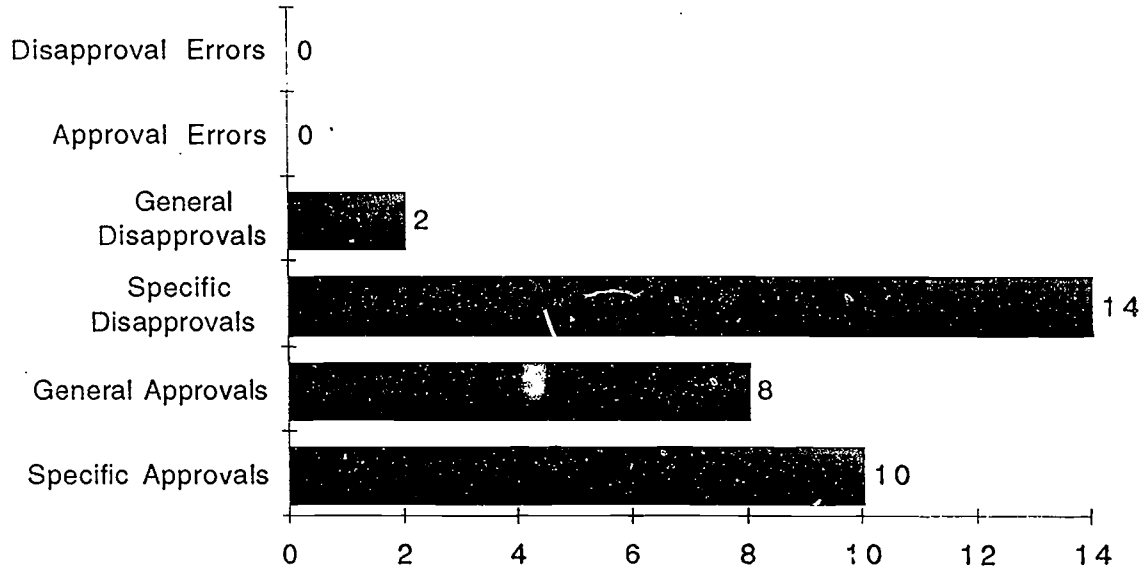


Figure 4
Ineffective Percentage Feedback Categories

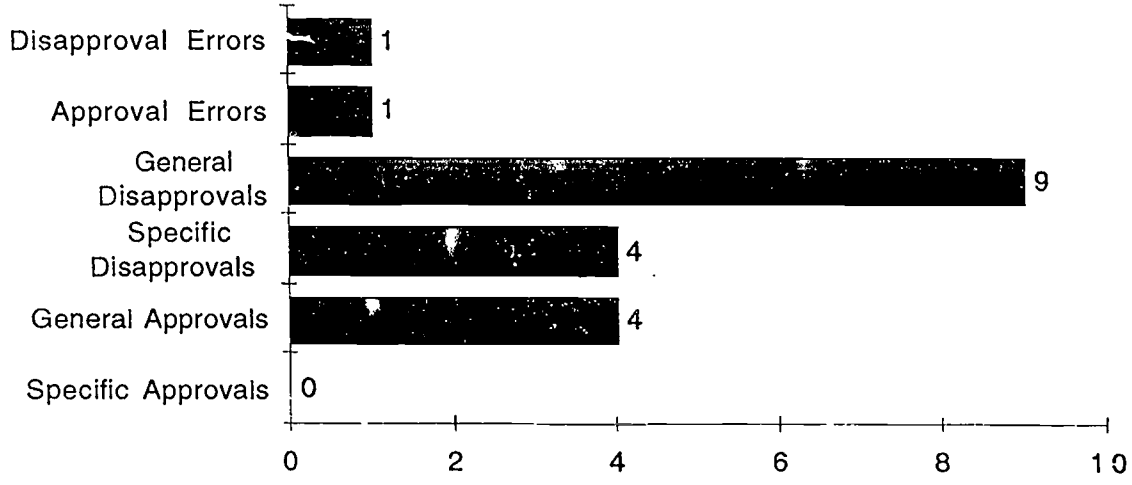


Figure 5
Effective Percentage Total Time for Student Behaviors

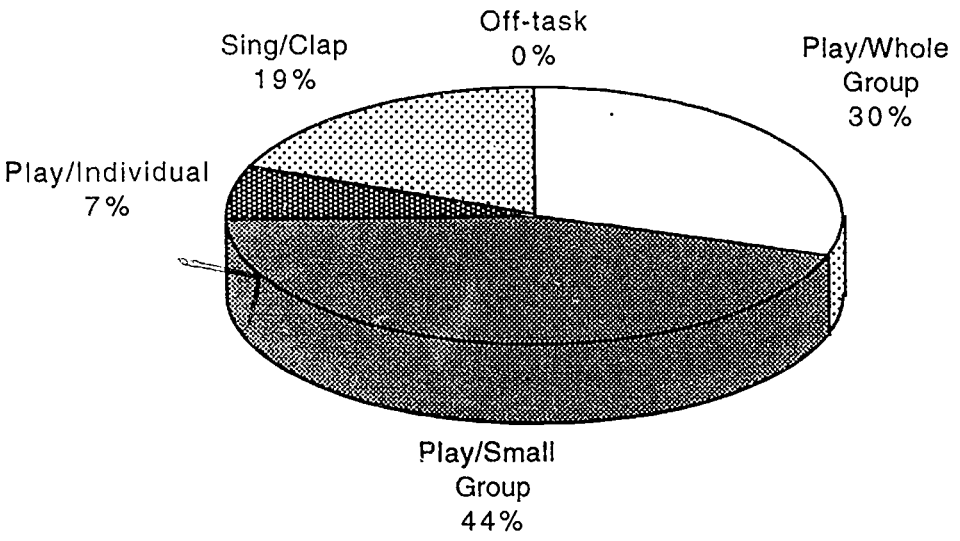
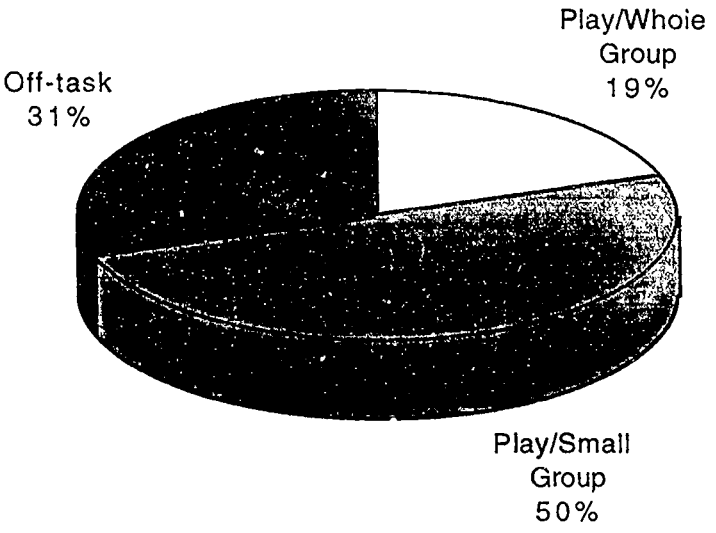


Figure 6
Ineffective Percentage of Time Use by Students



Graphic comparisons were made of selected variables for all subjects.

Figure 7
Off-Task Percentage Time for All Lesson Segments by Rank

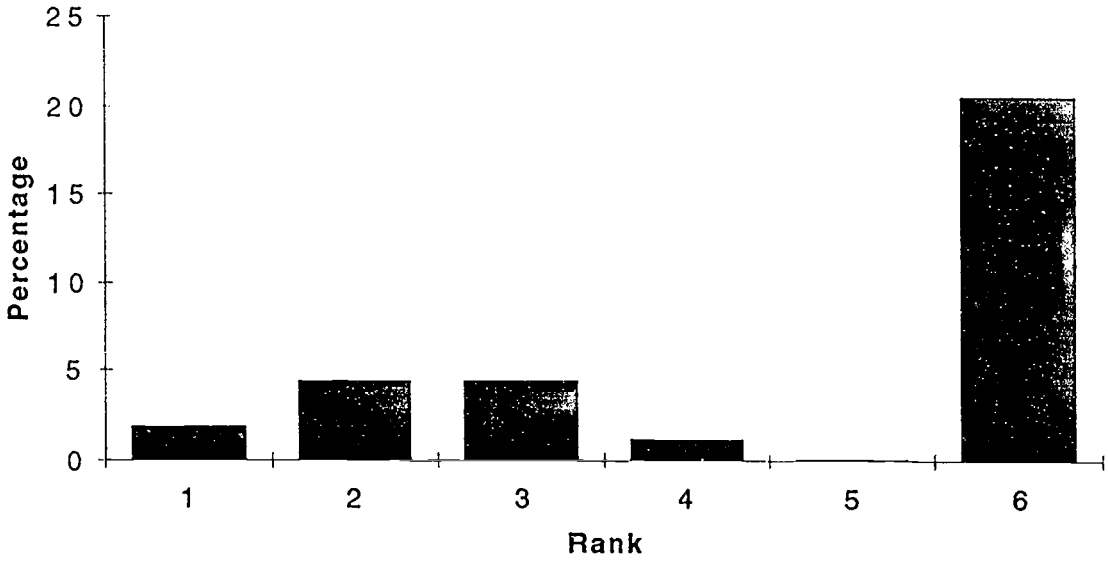


Figure 8
Specific Task Direction Frequency

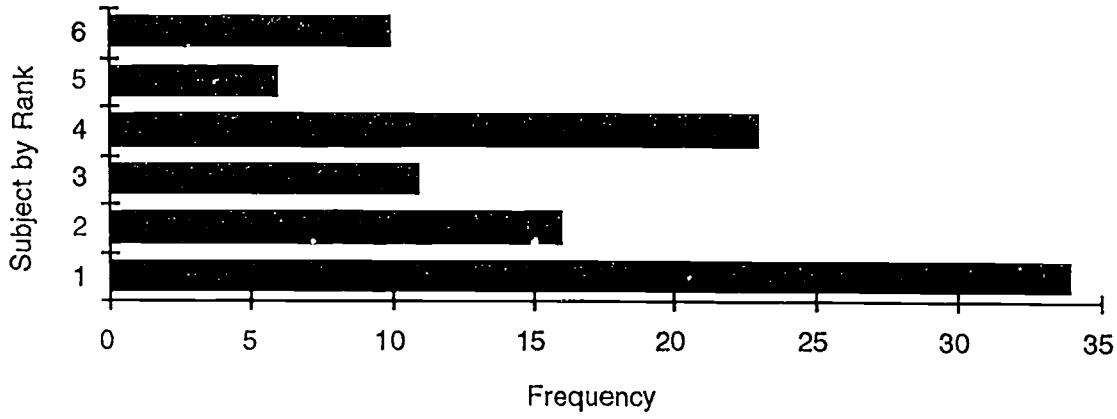


Figure 9
Frequency of Sing/Clap Modeling

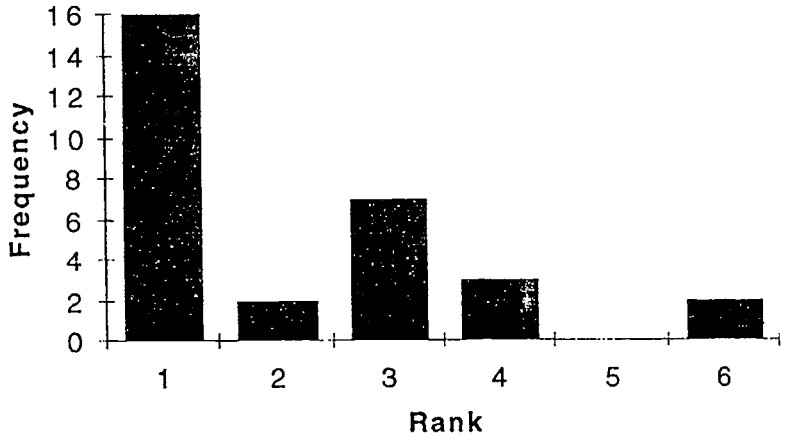


Figure 10
Approvals Comparison

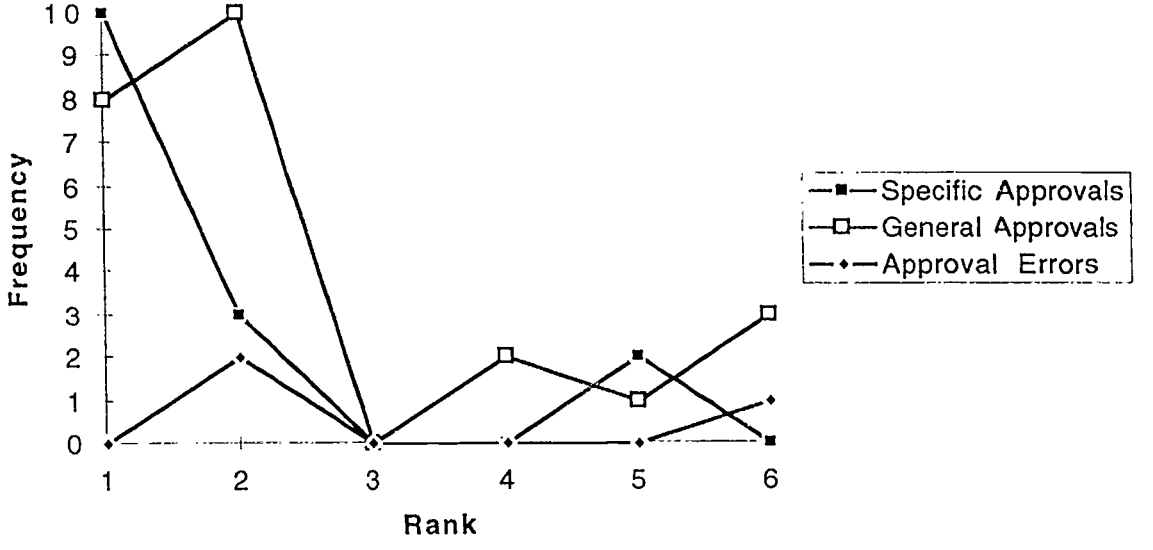


Figure 11
Comparison of Specific Disapprovals

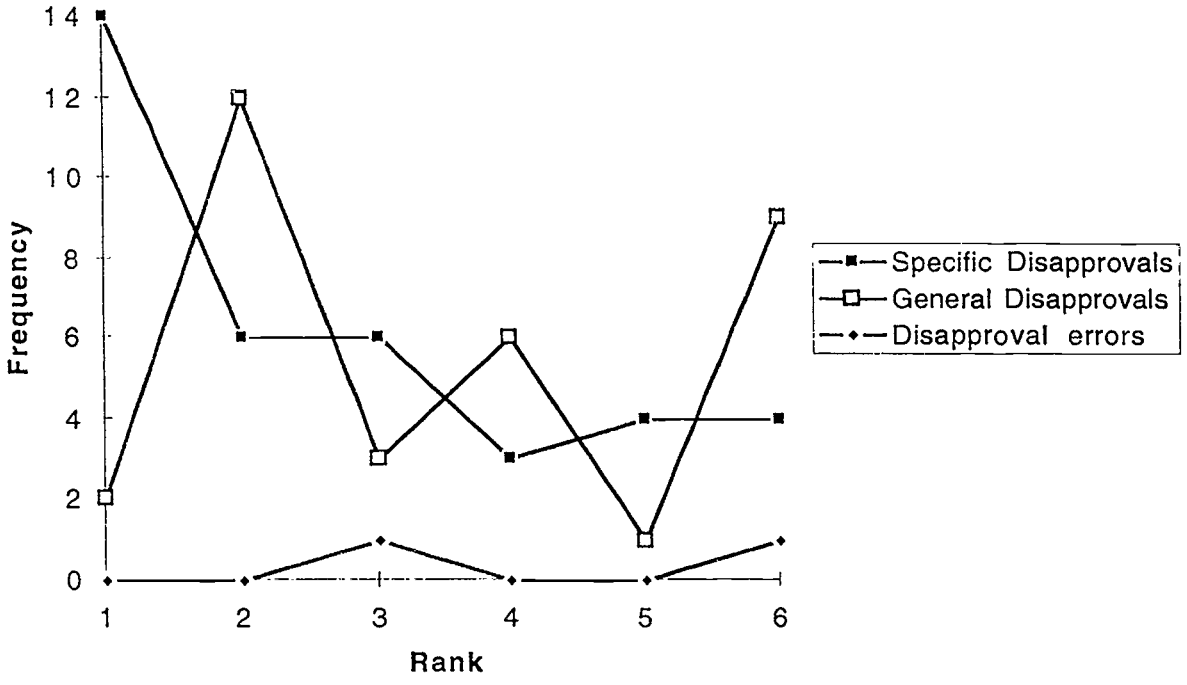


Figure 12
Comparison of Mean Playing Time

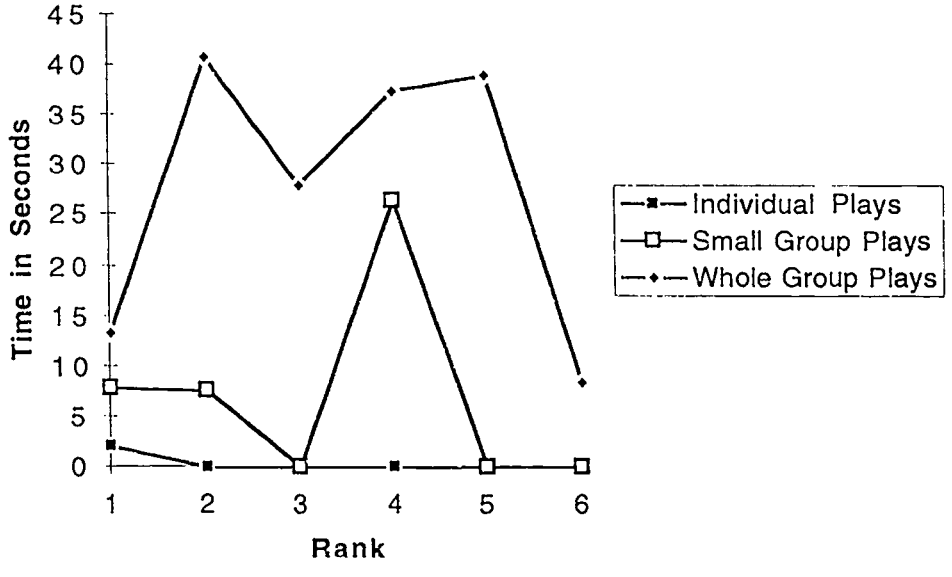
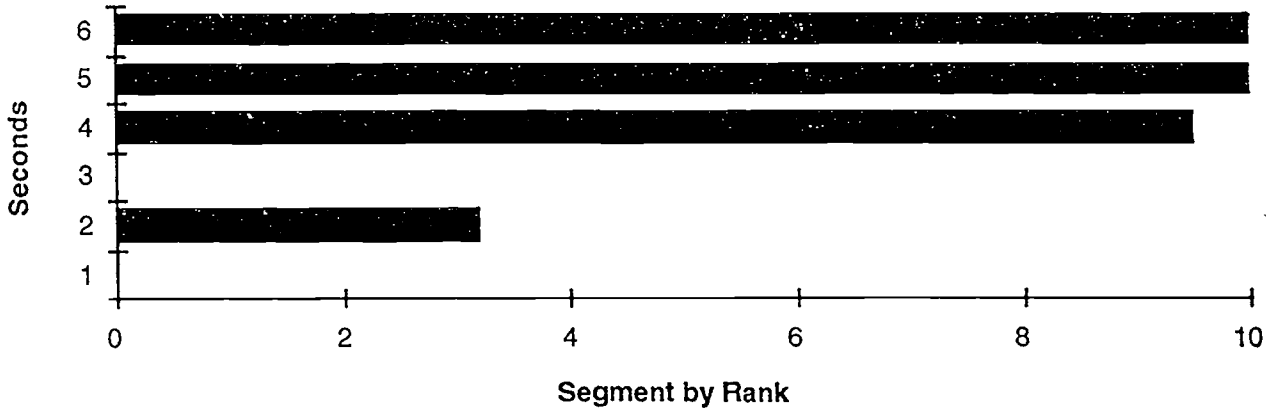


Figure 13
Mean Student Question/Answer Time



Conclusion

Testing the Method With Large-Group Instruction

This exploratory study has demonstrated that a systematic observation method using the new SCRIBE software is useful and generally manageable for gathering information on student-teacher interactions. The software was convenient particularly for recording the frequencies of behaviors and for summarizing the descriptive statistics for samples. The printout of behaviors in sequence along with the summaries of frequencies, durations, and rates was invaluable in pinpointing problem areas not apparent from the summaries. The timing of behaviors in sequence has received little attention in the past but merits further study.

In spite of the many advantages of the software, applications to large-group instruction need fine-tuning. Simultaneously occurring behaviors can be recorded only on multiple passes, creating sequences that are somewhat out of phase. This quirk was less problematic for individual instruction, where teacher and student behaviors tended to follow in turn. Behavior categories such as student performance and the durations of task directions that have been related to instructional effectiveness in studies of private instruction (Siebenaler, 1992; Younger Flores, 1989) were highly problematic in large-group settings. In many cases, it was difficult to

determine if the task was executed successfully because of the varying degrees of accuracy displayed by students. Furthermore, the teacher's intentions often were unclear, for example, "Try again at B," without stating any purpose for the repetition or target outcome.

A similar difficulty occurred in trying to determine the measures of task direction. Without a script or lesson plan, determining what (if anything) the teacher had in mind was virtually impossible, although the teacher's lack of direction may indicate a more serious problem, such as lack of planning. Perhaps, the addition of a category indicating uncertainty would help. Finding more accurate ways of relating teacher behaviors to the quality of student performance will be necessary for establishing truly valid measures of instructional effectiveness, especially for less traditional teaching strategies.

Comparisons Among Methods

Although quantitative determinations of the exact magnitude of the relationship between systematic observation data and global evaluations were beyond the scope of this small pilot study, the evidence clearly favors further investigation. Experts were highly reliable in selecting both effective and ineffective sequences. These results are somewhat surprising because of the use of experts within and from outside the field of music education. The summary and sequence data clearly support the qualitative differences expressed by the judges (see Figures 1-6). In the effective sample, the subject used specific task direction with nearly three times the frequency of the two lowest ranking samples (see Figure 8). The rate of sing/clap modeling was significantly higher. These findings concur with those of Sang (1982), who found clear directions indicative of effective instruction. The perception that less effective teachers somehow were not as involved or were occupied with issues outside of the learning sequence generally was supported in the observations (see Figure 7). With the exception of the second least effective sample, less time was wasted by the more effective subjects. The case of the fifth ranked subject is interesting. Though this subject spent the least amount of time off-task, the subject used too little corrective feedback and failed to properly isolate trouble spots. Optimizing instructional time during rehearsals also was supported as an indication of effective ensemble instruction in studies by Witt (1986) and Price (1983).

Feedback categories were somewhat limited in usefulness by problems with the student performance and task direction measures, but were very useful for clarifying ambiguities in ratings for the middle ranked samples. One subject was puzzling to the judges because in spite of many outstanding qualities, both personal and technical, judges had difficulty describing why the subject did not receive a higher ranking. Rigidity and reactivity rather than proactivity were cited as reasons. Examination with the software revealed a total absence of specific approvals

coupled with the lowest number of specific task directions. Thus, in spite of an appealing personality, the lack of guidance could prove very frustrating to students if allowed to continue. This is a clear example of how systematic observation data could prescribe a simple remedy to a problem that should easily be correctable by an otherwise promising director.

A high percentage of approvals is not the answer to the ills of ineffective instruction. In fact, the highest ranked director used the greatest number of specific disapprovals. Discounting errors in feedback, the most and least effective samples differed substantially in the ratio of specific to general feedback and in another remarkable way—the timing of disapprovals. The process of shaping skill behaviors by nature temporarily decreases the student's sense of efficacy as she or he is presented with a seemingly insurmountable difficulty to be overcome under the watchful eyes of teacher and peers. If the teacher isn't careful to help rebuild what has been shaken, the student may feel inadequate. The effective example displayed many corrections but ended on a positive note. The ensemble was told how all of the hard work on rhythmic precision and uniform articulation resulted in the band sounding like one instrument. The least effective sample went from somewhat approving to downright insulting, and the director appeared to give up and jump to a new activity in response to the breakdowns in musical performance.

Although conductors cannot solve all of the uncovered difficulties within 10 minutes, discouraging rehearsals that become the rule may lead to attrition. Music should be challenging, but fun. Students are unlikely to pursue an activity that makes them feel inept in front of peers. Teachers are equally likely to flee from situations that seem headed on a downward spiral. Going back to the evaluator comment of "poor choice of music" to see if the piece was too difficult (or just too boring), and examining the sequence of directive and feedback behaviors could help the teacher structure the rehearsal for more successful learning outcomes.

Although the means and standard deviations of student playing times were highly significant in preliminary findings by Robert A. Duke (personal communication, March 1995) at the University of Texas, the results of the present study were inconclusive, except that very long mean playing times with little variance was obvious in the case of the fifth ranked subject. In fact, the data support the judges' comments that the subject "was not teaching enough," although it was impossible to tell from the sample if uninterrupted playing time was the goal for that lesson.

Modifications To Accommodate Diverse Learners

Although no extraordinary modifications were used to accommodate diverse learners, commentaries from both music educators and general education evaluators posed some new avenues for exploration. Although the practice of writing examples on the board was not related

to improved playing nor to ranking, seeking peer input was related to higher ranking for all but the top-ranked subject. The authoritarian role of the conductor has been a given fact. He or she is deemed the expert and seeking feedback from the members of the orchestra or band often is viewed as inefficient or ineffective, or both. No one knows if lack of systematic training in critical listening skills is responsible, at least in part, for the time wasted during individual practice sessions. Although question and answer sessions correlated strongly with lower subject rank in this study and with poor student attitude in the previously mentioned studies of Siebenaler (1992) and Bolte (1994), students' solicitation of feedback is deemed highly desirable by experts of teacher assessment. Only further study with better dependent measures for student progress in the short and long run would determine the ultimate usefulness of the shift from an authoritarian environment to a more democratic one. This shift in belief systems should be studied for the potential effect on musical performance, language development, and subjective measures of student attitude and self-efficacy before conclusions are drawn.

The present study was limited in scope and all findings should be regarded with proper caution. Without further study, automatic generalizations should be avoided. The usefulness of the study is to confirm the importance of examining multiple means of determining instructional effectiveness to identify patterns of behaviors characteristic of good teaching. Apparently, a complementary relationship exists between global evaluations and systematic observations, such that systematic observations may be used to clarify vague perceptions of teaching quality. Commentaries used to explain rankings may suggest new or better observation categories. Use of the two methods was enhanced by observing a sample of effective teaching, which was used as a model for comparing and contrasting less effective or differing teaching styles.

Continued fine tuning of the software should make the computer-assisted systematic observation method even more useful in the future. Plans are in progress to display outputs directly in graphic form. Simplifying the procedures will make systematic observation a more reasonable choice for assess ng teacher effectiveness and for teachers' self-study, and a more useful tool to researchers in teaching methodology and skill-based learning. Helping teachers to make playing in an ensemble a challenging but rewarding experience for all music students should be a goal of successful instructional evaluation.

References

Bloom, B. (Ed.) (1985). *Developing talent in young people*. New York: Ballantine Books.

Bolte, K. (1994). *Observation of chemistry instruction using the Kieler Observation Instrument.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Brophy, J., & Good, T. (1986). Teacher behavior and student achievement. In M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 340-370). New York: Macmillan.

Duke, R. A., & Madsen, C. K. (1991). Proactive versus reactive teaching: Focusing observation on specific aspects of instruction. *Bulletin of the Council for Research in Music Education, 108*, 1-14.

Duke, R. A., & Prickett, C. A. (1992). Evaluation of music instruction by musicians and nonmusicians assigned differential observation tasks. *Bulletin of the Council for Research in Music Education, 113*, 41-50.

Fitts, P. M., & Posner, M. I. (1967). *Human performance*. Belmont, CA: Wadsworth.

Flanders, N. A. (1964). *Interaction analysis in the classroom.* Ann Arbor, MI: University of Michigan School of Education.

Flanders, N. A. (1970). *Analyzing teacher behavior.* Reading, MA: Addison-Wesley.

Gagne, R. M. (1977). *The conditions of learning*. New York: Holt, Rinehart, and Winston.

Gagne, R. M., Briggs, L. J., & Wager, W. (1988). *Principles of instructional design.* New York: Holt, Rinehart, and Winston.

Lawrence, S. J. (1978). *Behavioral profile of the piano student.* Hewlett, NY: Workshop Music Teaching Publications.

National Evaluation Systems, Inc. (1994-95). *Information guide: Assessment of teaching skills, New York State Teacher Certification Examinations*. Amherst, MA: Author.

Obiakor, F. (1992). Embracing new special education strategies for African American students. *Exceptional Children, 59*(2), 104-106.

Price, H. (1983). The effect of conductor academic task presentation, conductor reinforcement, and ensemble practice on performers' musical achievement, attentiveness, and attitude. *Journal of Research in Music Education, 31,* 247-257.

Richard-Amato, P. (1988). *Making it happen: Interactions in the second language classroom.* New York: Longman.

Sang, R. (1982). *Modified path analysis of a skills-based instructional effectiveness model for beginning teachers in instrumental music education.* Doctoral dissertation, the University of Michigan.

Shaver, J. P. (1983, October). The verification of independent variables in teaching method research. *Education Researcher,* 3-9.

Siebenaler, D. (1992). *Analysis of the student-teacher interactions in the lessons of children and adults.* Doctoral treatise, the University of Texas at Austin.

Swank Paul R. et al. (1989). Sensitivity of classroom observation systems: Measuring teacher effectiveness. *Journal of Experimental Education, 57*(2), 171-86.

Tharp, R., & Gallimore, R. (1988). *Rousing minds to life.* Cambridge, England: Cambridge University Press.

Witt, A. C. (1986). Use of class time and student attentiveness in secondary instrumental music rehearsals. *Journal of Research in Music Education, 33,* 34-42.

Wright, J. V. (1995). Multicultural issues and attention deficit disorders. *Learning Disabilities Research and Practice, 103*(3), 153-159.

Younger Flores, K. (1989). *Systematic observation of Suzuki method violin instruction.* Unpublished master's project, the University of Texas at Austin.

Appendix A

SCRIBE Output for Highest-Ranked Excerpt

Test Chronology

Test Started: 1/1/4 12:58:52 AM

No	Start	Subject	Behavior	Duration
1	0:00:03	Teacher	G	0:04 *
2	0:00:07	Teacher	OT	0:18 ****
3	0:00:23	Students	p	1:00 **********
4	0:00:25	Teacher	G	0:08 **
5	0:00:33	Teacher	T	0:01 *
6	0:01:32	Students	+	0:06 **
7	0:01:38	Progress	->	0:02 *
8	0:01:47	Teacher	a	0:02 *
9	0:01:49	Teacher	A	0:03 *
10	0:01:53	Teacher	A	0:02 *
11	0:01:55	Teacher	T	0:18 ****
12	0:01:57	Students	p	0:02 *
13	0:02:01	Students	-	0:01 *
14	0:02:02	Progress	<-	0:03 *
15	0:02:05	Students	p	0:06 **
16	0:02:11	Students	-	0:01 *
17	0:02:12	Progress	/	0:03 *
18	0:02:13	Teacher	G	0:03 *
19	0:02:15	Students	p	0:03 *
20	0:02:24	Teacher	D	0:02 *
21	0:02:26	Teacher	G	0:01 *
22	0:02:28	Students	s/c	0:05 *
23	0:02:31	Teacher	D	0:02 *
24	0:02:33	Students	+	0:06 **
25	0:02:34	Teacher	G	0:02 *
26	0:02:39	Progress	/	0:04 *
27	0:02:43	Students	s/c	0:03 *
28	0:02:44	Teacher	T	0:01 *
29	0:02:46	Students	+	0:01 *
30	0:02:47	Progress	/	0:02 *
31	0:02:47	Teacher	M	0:01 *
32	0:02:49	Students	s/c	0:04 *
33	0:02:52	Teacher	D	0:01 *
34	0:02:53	Progress	->	0:03 *
35	0:02:55	Teacher	M	0:01 *
36	0:02:57	Teacher	M	0:02 *
37	0:02:59	Teacher	d	0:01 *
38	0:03:02	Students	s/c	0:04 *
39	0:03:06	Students	-	0:03 *
40	0:03:06	Teacher	T	0:05 *
41	0:03:09	Progress	/	0:07 **
42	0:03:11	Teacher	G	0:02 *
43	0:03:13	Teacher	T	0:01 *
44	0:03:14	Teacher	G	0:02 *
45	0:03:16	Teacher	M	0:01 *
46	0:03:16	Students	s/c	0:04 *
47	0:03:18	Teacher	G	0:03 *
48	0:03:20	Students	+	0:02 *
49	0:03:25	Progress	->	0:05 *
50	0:03:26	Teacher	D	0:04 *
51	0:03:31	Teacher	M	0:01 *
52	0:03:32	Teacher	T	0:01 *
53	0:03:34	Teacher	G	0:01 *

54	0:03:35	Students	p	0:05 *
55	0:03:40	Students	-	0:04 *
56	0:03:44	Progress	/	0:10 **
57	0:03:46	Teacher	T	0:08 **
58	0:03:54	Teacher	G	0:02 *
59	0:03:58	Students	p	0:04 *
60	0:04:01	Teacher	M	0:04 *
61	0:04:05	Students	-	0:02 *
62	0:04:07	Progress	/	0:02 *
63	0:04:08	Teacher	D	0:02 *
64	0:04:09	Students	-	0:05 *
65	0:04:10	Teacher	M	0:04 *
66	0:04:14	Progress	<-	0:07 **
67	0:04:14	Teacher	T	0:02 *
68	0:04:16	Teacher	G	0:01 *
69	0:04:21	Students	+	0:04 *
70	0:04:25	Progress	/	0:04 *
71	0:04:26	Teacher	G	0:02 *
72	0:04:29	Students	p	0:06 **
73	0:04:32	Teacher	T	0:02 *
74	0:04:34	Teacher	M	0:02 *
75	0:04:35	Students	P/SG	0:05 *
76	0:04:36	Teacher	T	0:01 *
77	0:04:38	Teacher	G	0:01 *
78	0:04:40	Students	+	0:07 **
79	0:04:47	Students	PI	0:02 *
80	0:04:48	Teacher	M	0:05 *
81	0:04:49	Students	+	0:02 *
82	0:04:51	Progress	/	0:01 *
83	0:04:52	Students	PI	0:02 *
84	0:04:54	Students	-	0:01 *
85	0:04:55	Students	PI	0:02 *
86	0:04:55	Teacher	M	0:13 ***
87	0:04:57	Students	+	0:06 **
88	0:05:03	Progress	->	0:01 *
89	0:05:04	Students	P/SG	0:04 *
90	0:05:08	Teacher	G	0:01 *
91	0:05:08	Students	-	0:02 *
92	0:05:10	Progress	/	0:01 *
93	0:05:11	Teacher	a	0:02 *
94	0:05:11	Students	P/SG	0:04 *
95	0:05:13	Teacher	G	0:01 *
96	0:05:15	Students	+	0:03 *
97	0:05:16	Teacher	a	0:03 *
98	0:05:18	Progress	->	0:04 *
99	0:05:19	Teacher	T	0:01 *
100	0:05:21	Teacher	G	0:01 *
101	0:05:22	Students	P/SG	0:03 *
102	0:05:24	Teacher	A	0:01 *
103	0:05:28	Progress	/	0:03 *
104	0:05:30	Teacher	D	0:03 *
105	0:05:31	Students	P/SG	0:02 *
106	0:05:33	Teacher	G	0:12 ***
107	0:05:33	Students	-	0:05 *
108	0:05:38	Progress	->	0:02 *
109	0:05:40	Students	P/SG	0:05 *
110	0:05:45	Students	-	0:03 *
111	0:05:45	Teacher	G	0:01 *
112	0:05:48	Teacher	G	0:02 *
113	0:05:48	Progress	->	0:03 *

114	0:05:50	Teacher	T	0:01 *
115	0:05:51	Students	P/SG	0:03 *
116	0:05:51	Teacher	G	0:01 *
117	0:05:54	Students	+	0:03 *
118	0:05:56	Teacher	G	0:06 **
119	0:05:57	Progress	<-	0:02 *
120	0:05:59	Students	P/SG	0:05 *
121	0:06:04	Students	-	0:02 *
122	0:06:05	Teacher	a	0:01 *
123	0:06:06	Progress	/	0:03 *
124	0:06:08	Teacher	G	0:01 *
125	0:06:09	Students	P/SG	0:03 *
126	0:06:14	Teacher	A	0:01 *
127	0:06:16	Students	+	0:06 **
128	0:06:17	Teacher	G	0:02 *
129	0:06:22	Progress	/	0:01 *
130	0:06:23	Students	-	0:04 *
131	0:06:24	Teacher	T	0:02 *
132	0:06:26	Teacher	G	0:03 *
133	0:06:27	Students	P/SG	0:04 *
134	0:06:31	Students	-	0:03 *
135	0:06:33	Teacher	G	0:05 *
136	0:06:34	Progress	/	0:04 *
137	0:06:38	Students	s/c	0:05 *
138	0:06:42	Teacher	M	0:06 **
139	0:06:44	Students	-	0:02 *
140	0:06:46	Progress	/	0:05 *
141	0:06:48	Teacher	G	0:00
142	0:06:54	Students	s/c	0:02 *
143	0:06:54	Teacher	T	0:04 *
144	0:06:56	Students	+	0:02 *
145	0:06:58	Progress	/	0:00
146	0:07:03	Students	s/c	0:04 *
147	0:07:05	Teacher	T	0:02 *
148	0:07:07	Students	-	0:03 *
149	0:07:08	Teacher	M	0:03 *
150	0:07:10	Progress	->	0:01 *
151	0:07:11	Students	p	0:05 *
152	0:07:15	Teacher	a	0:01 *
153	0:07:16	Teacher	G	0:01 *
154	0:07:16	Students	+	0:02 *
155	0:07:19	Teacher	D	0:01 *
156	0:07:21	Progress	->	0:02 *
157	0:07:22	Teacher	G	0:01 *
158	0:07:28	Teacher	G	0:02 *
159	0:07:31	Students	p	0:10 **
160	0:07:41	Teacher	T	0:09 **
161	0:07:44	Students	+	0:02 *
162	0:07:46	Progress	->	0:03 *
163	0:07:52	Teacher	G	0:02 *
164	0:07:56	Students	p	0:20 ****
165	0:08:02	Teacher	G	0:13 ***
166	0:08:16	Students	+	0:02 *
167	0:08:23	Progress	<-	0:01 *
168	0:08:31	Students	P/SG	0:14 ***
169	0:08:38	Teacher	a	0:02 *
170	0:08:42	Teacher	G	0:06 **
171	0:08:47	Students	-	0:03 *
172	0:08:48	Teacher	M	0:02 *
173	0:08:50	Progress	<-	0:04 *

174	0:08:53	Teacher	G	0:01 *
175	0:09:06	Students	P/SG	0:05 *
176	0:09:10	Teacher	D	0:11 ***
177	0:09:11	Students	-	0:04 *
178	0:09:15	Progress	/	0:02 *
179	0:09:17	Students	P/SG	0:20 ****
180	0:09:21	Teacher	T	0:03 *
181	0:09:24	Teacher	M	0:02 *
182	0:09:26	Teacher	G	0:05 *
183	0:09:31	Teacher	d	0:02 *
184	0:09:33	Teacher	D	0:02 *
185	0:09:35	Teacher	M	0:01 *
186	0:09:36	Teacher	T	0:02 *
187	0:09:37	Students	+	0:03 *
188	0:09:39	Teacher	T	0:01 *
189	0:09:40	Progress	->	0:03 *
190	0:09:40	Teacher	G	0:04 *
191	0:09:44	Teacher	a	0:02 *
192	0:09:46	Teacher	G	0:04 *
193	0:09:55	Students	P/SG	0:17 ****
194	0:09:57	Teacher	A	0:02 *
195	0:10:02	Teacher	G	0:12 ***
196	0:10:22	Students	+	0:01 *
197	0:10:37	Progress	/	0:03 *
198	0:10:39	Teacher	D	0:07 **
199	0:10:40	Students	P/SG	0:17 ****
200	0:10:46	Teacher	T	0:03 *
201	0:10:49	Teacher	G	0:03 *
202	0:10:52	Teacher	T	0:04 *
203	0:10:56	Teacher	G	0:11 ***
204	0:11:00	Students	+	0:04 *
205	0:11:07	Teacher	a	0:11 ***
206	0:11:14	Progress	/	0:05 *
207	0:11:18	Teacher	A	0:12 ***
208	0:11:19	Students	P/SG	0:16 ****
209	0:11:30	Teacher	T	0:02 *
210	0:11:32	Teacher	G	0:01 *
211	0:11:33	Teacher	A	0:03 *
212	0:11:36	Teacher	G	0:04 *
213	0:11:39	Students	+	0:08 **
214	0:11:47	Teacher	T	0:01 *
215	0:11:47	Progress	->	0:03 *
216	0:11:50	Students	P/SG	0:03 *
217	0:11:53	Students	-	0:02 *
218	0:12:00	Teacher	T	0:02 *
219	0:12:00	Progress	/	0:05 *
220	0:12:05	Teacher	G	0:04 *
221	0:12:11	Students	P/SG	0:10 **
222	0:12:14	Teacher	D	0:10 **
223	0:12:24	Students	+	0:04 *
224	0:12:26	Teacher	T	0:00
225	0:12:27	Teacher	T	0:01 *
226	0:12:29	Teacher	T	0:01 *
227	0:12:30	Teacher	G	0:02 *
228	0:12:35	Students	-	0:02 *
229	0:12:37	Progress	/	0:07 **
230	0:12:44	Students	P/SG	0:11 ***
231	0:12:47	Teacher	MT	0:01 *
232	0:12:51	Teacher	D	0:05 *
233	0:12:55	Students	+	0:02 *

234	0:12:56	Teacher	T	0:02 *
235	0:12:57	Progress	->	0:12 ***
236	0:12:58	Teacher	M	0:02 *
237	0:13:00	Teacher	A	0:05 *
238	0:13:09	Teacher	T	0:02 *
239	0:13:09	Students	p	0:03 *
240	0:13:11	Teacher	T	0:01 *
241	0:13:12	Students	-	0:04 *
242	0:13:15	Teacher	A	0:02 *
243	0:13:16	Progress	/	0:02 *
244	0:13:17	Teacher	G	0:01 *
245	0:13:18	Students	p	0:27 ******
246	0:13:21	Teacher	T	0:02 *
247	0:13:26	Teacher	G	0:01 *
248	0:13:35	Teacher	D	0:03 *
249	0:13:38	Teacher	G	0:01 *
250	0:13:45	Students	-	0:07 **
251	0:13:52	Progress	/	0:06 **
252	0:13:59	Teacher	T	0:03 *
253	0:14:03	Students	p	0:22 *****
254	0:14:06	Teacher	D	0:05 *
255	0:14:13	Teacher	T	0:01 *
256	0:14:17	Teacher	T	0:02 *
257	0:14:20	Teacher	G	0:02 *
258	0:14:25	Students	+	0:03 *
259	0:14:41	Teacher	A	0:05 *

Test Stopped: 1/1/4 1:29:25 AM

Test Summary

Test Time: 0:14:46
Number of Passes: 2

	Frequency	Time (Min:Sec)	Time (%)	Mean (Min:Sec)	Standard Deviation	Rate (#/min)
Teacher						
T	34	1:32	10.4	0:02.7	3.259	2.3
G	47	2:29	16.8	0:03.2	3.158	3.2
M	16	0:50	5.6	0:03.1	2.955	1.1
MT	1	0:01	0.1	0:01.0	0.000	0.1
A	10	0:36	4.1	0:03.6	3.105	0.7
a	8	0:24	2.7	0:03.0	3.082	0.5
D	14	0:58	6.5	0:04.1	3.067	0.9
OT	1	0:18	2.0	0:18.0	0.000	0.1
d	2	0:03	0.3	0:01.5	0.500	0.1
Students						
p	13	2:53	19.5	0:13.3	15.617	0.9
s/c	8	0:31	3.5	0:03.9	0.927	0.5
P/SG	19	2:31	17.0	0:07.9	5.799	1.3
PI	3	0:06	0.7	0:02.0	0.000	0.2
+	22	1:19	8.9	0:03.6	1.992	1.5
-	21	1:03	7.1	0:03.0	1.480	1.4

Appendix B

SCRIBE Output for Lowest-Ranked Excerpt

Test Chronology

Test Started: 1/1/4 8:25:56 AM

No	Start	Subject	Behavior	Duration
1	0:00:02	teacher	G	0:15 ***
2	0:00:13	Students	play	0:19 ****
3	0:00:26	Progress	/	0:02 *
4	0:00:27	teacher	SD	0:01 *
5	0:00:32	Students	+	0:02 *
6	0:00:33	Progress	<-	0:08 **
7	0:00:34	teacher	G	0:01 *
8	0:00:39	Students	play	0:05 *
9	0:00:44	Students	+	0:02 *
10	0:00:46	Progress	X	0:07 **
11	0:00:49	Students	SG	0:03 *
12	0:00:50	teacher	D	0:02 *
13	0:00:52	Students	-	0:02 *
14	0:00:55	teacher	dm	0:01 *
15	0:00:56	Progress	<-	0:05 *
16	0:00:59	teacher	G	0:01 *
17	0:00:59	Students	OT	0:10 **
18	0:01:06	teacher	G	0:02 *
19	0:01:07	Progress	/	0:03 *
20	0:01:12	Progress	/	0:05 *
21	0:01:12	teacher	G	0:03 *
22	0:01:14	Students	play	0:04 *
23	0:01:15	teacher	SD	0:01 *
24	0:01:20	Students	SG	0:08 **
25	0:01:23	Progress	/	0:01 *
26	0:01:23	teacher	G	0:02 *
27	0:01:26	Progress	->	0:19 ****
28	0:01:28	Students	+	0:04 *
29	0:01:32	Students	play	0:02 *
30	0:01:34	Students	-	0:02 *
31	0:01:39	Students	play	0:16 ****
32	0:01:42	teacher	OT	0:31 *******
33	0:01:45	Progress	X	0:25 *****
34	0:01:55	Students	+	0:02 *
35	0:01:59	Students	OT	0:30 ******
36	0:02:10	Progress	X	0:05 *
37	0:02:13	teacher	a	0:02 *
38	0:02:15	teacher	G	0:02 *
39	0:02:16	Progress	<-	0:03 *
40	0:02:18	teacher	G	0:01 *
41	0:02:27	Progress	->	0:03 *
42	0:02:28	teacher	G	0:01 *
43	0:02:31	Students	SG	0:12 ***
44	0:02:43	Students	+	0:02 *
45	0:02:44	teacher	G	0:01 *
46	0:02:45	Students	SG	0:13 ***
47	0:02:47	Progress	/	0:01 *
48	0:02:49	teacher	d	0:02 *
49	0:02:50	Progress	<-	0:01 *
50	0:02:51	teacher	d	0:01 *
51	0:02:52	Progress	X	0:16 ****
52	0:02:54	teacher	D	0:02 *
53	0:02:58	teacher	OT	0:02 *

54	0:03:00	teacher	d	0:03 *
55	0:03:02	Students	-	0:02 *
56	0:03:03	teacher	D	0:01 *
57	0:03:04	Students	SG	0:02 *
58	0:03:04	teacher	OT	0:04 *
59	0:03:06	Students	-	0:02 *
60	0:03:08	Students	OT	0:15 ***
61	0:03:08	teacher	d	0:01 *
62	0:03:08	Progress	<-	0:03 *
63	0:03:09	teacher	G	0:01 *
64	0:03:11	Progress	/	0:02 *
65	0:03:13	teacher	G	0:01 *
66	0:03:16	teacher	d	0:02 *
67	0:03:18	Progress	/	0:03 *
68	0:03:21	teacher	G	0:01 *
69	0:03:23	Students	SG	0:06 **
70	0:03:24	teacher	a	0:01 *
71	0:03:26	Progress	/	0:02 *
72	0:03:27	teacher	G	0:02 *
73	0:03:29	teacher	SD	0:02 *
74	0:03:29	Students	-	0:04 *
75	0:03:33	teacher	d	0:02 *
76	0:03:34	Students	OT	0:08 **
77	0:03:37	teacher	OT	0:04 *
78	0:03:38	Progress	X	0:03 *
79	0:03:42	Progress	<-	0:02 *
80	0:03:42	Students	SG	0:05 *
81	0:03:42	teacher	SD	0:01 *
82	0:03:44	teacher	G	0:02 *
83	0:03:47	Students	-	0:04 *
84	0:03:49	Progress	->	0:05 *
85	0:03:51	Students	OT	0:07 **
86	0:03:51	teacher	G	0:01 *
87	0:03:56	teacher	SD	0:01 *
88	0:03:58	Progress	/	0:12 ***
89	0:03:58	Students	SG	0:15 ***
90	0:04:01	teacher	G	0:02 *
91	0:04:04	teacher	S/C	0:01 *
92	0:04:07	teacher	G	0:02 *
93	0:04:10	Progress	<-	0:06 **
94	0:04:12	teacher	d	0:02 *
95	0:04:13	Students	+	0:02 *
96	0:04:14	teacher	SD	0:01 *
97	0:04:16	teacher	G	0:02 *
98	0:04:17	Progress	<-	0:04 *
99	0:04:18	teacher	d	0:02 *
100	0:04:18	Students	OT	0:04 *
101	0:04:20	teacher	SD	0:02 *
102	0:04:22	Students	SG	0:02 *
103	0:04:22	Progress	->	0:04 *
104	0:04:23	teacher	G	0:04 *
105	0:04:24	Students	-	0:05 *
106	0:04:28	Progress	->	0:20 ****
107	0:04:29	Students	SG	0:03 *
108	0:04:29	teacher	a	0:01 *
109	0:04:32	Students	-	0:02 *
110	0:04:33	teacher	G	0:02 *
111	0:04:34	Students	SG	0:08 **
112	0:04:41	teacher	a	0:01 *
113	0:04:42	Students	+	0:02 *

114	0:04:44	Students	SG	0:12	***
115	0:04:46	teacher	G	0:03	*
116	0:04:49	Progress	->	0:09	**
117	0:04:57	Students	+	0:04	*
118	0:04:59	Progress	->	0:03	*
119	0:05:01	Students	SG	0:13	***
120	0:05:02	Progress	->	0:01	*
121	0:05:04	teacher	G	0:01	*
122	0:05:06	Progress	/	0:02	*
123	0:05:07	teacher	G	0:01	*
124	0:05:09	Progress	<-	0:09	**
125	0:05:14	Students	+	0:05	*
126	0:05:19	Students	OT	0:03	*
127	0:05:22	Progress	X	0:09	**
128	0:05:22	Students	SG	0:02	*
129	0:05:24	Students	play	0:13	***
130	0:05:26	teacher	G	0:02	*
131	0:05:34	teacher	OT	0:04	*
132	0:05:37	Students	+	0:03	*
133	0:05:38	teacher	G	0:02	*
134	0:05:45	Students	OT	0:09	**
135	0:05:54	Students	play	0:48	**********
136	0:06:30	teacher	G	0:03	*
137	0:06:30	Progress	/	0:09	**
138	0:06:35	teacher	G	0:01	*
139	0:06:37	teacher	G	0:02	*
140	0:06:42	Students	-	0:02	*
141	0:06:50	Students	OT	0:02	*
142	0:06:52	Students	SG	0:15	***
143	0:06:53	teacher	am	0:03	*
144	0:06:54	Progress	X	0:14	***
145	0:06:56	teacher	OT	0:07	**
146	0:07:03	teacher	G	0:02	*
147	0:07:05	teacher	OT	0:03	*
148	0:07:07	Students	-	0:03	*
149	0:07:08	teacher	SD	0:05	*
150	0:07:09	Progress	/	0:13	***
151	0:07:10	Students	OT	0:08	**
152	0:07:13	teacher	S/C	0:03	*
153	0:07:16	teacher	G	0:04	*
154	0:07:22	Progress	/	0:02	*
155	0:07:24	Progress	/	0:03	*
156	0:07:26	teacher	G	0:01	*
157	0:07:27	teacher	SD	0:02	*
158	0:07:28	Progress	/	0:06	**
159	0:07:29	teacher	G	0:09	**
160	0:07:33	Students	SG	0:05	*
161	0:07:35	Progress	->	0:28	******
162	0:07:38	Students	+	0:02	*
163	0:07:52	Students	SG	0:26	******
164	0:08:05	teacher	d	0:01	*
165	0:08:08	teacher	D	0:02	*
166	0:08:10	Progress	<-	0:02	*
167	0:08:12	teacher	OT	0:04	*
168	0:08:16	teacher	G	0:01	*
169	0:08:17	teacher	SD	0:01	*
170	0:08:18	teacher	G	0:06	**
171	0:08:18	Students	-	0:01	*
172	0:08:19	Students	OT	0:11	***
173	0:08:30	Students	SG	0:03	*

174 0:08:35 Students + 0:08 **

Test Stopped: 1/1/4 8:55:12 AM

Test Summary

Test Time: 0:08:43
Number of Passes: 3

	Frequency	Time (Min:Sec)	Time (%)	Mean (Min:Sec)	Standard Deviation	Rate (#/min)
teacher						
SD	10	0:17	3.3	0:01.7	1.187	1.1
G	35	1:27	16.6	0:02.5	2.677	4.0
S/C	2	0:04	0.8	0:02.0	1.000	0.2
OT	8	0:59	11.3	0:07.4	9.027	0.9
a	4	0:05	1.0	0:01.2	0.433	0.5
D	4	0:07	1.3	0:01.8	0.433	0.5
d	9	0:16	3.1	0:01.8	0.629	1.0
am	1	0:03	0.6	0:03.0	0.000	0.1
dm	1	0:01	0.2	0:01.0	0.000	0.1
Students						
play	7	1:47	20.5	0:15.3	14.636	0.8
SG	18	2:33	29.3	0:08.5	6.256	2.1
OT	11	1:47	20.5	0:09.7	7.337	1.3
+	12	0:38	7.3	0:03.2	1.772	1.4
-	11	0:29	5.5	0:02.6	1.150	1.3
Progress						
->	9	1:32	17.6	0:10.2	9.102	1.0
<-	10	0:43	8.2	0:04.3	2.532	1.1
/	15	1:06	12.6	0:04.4	3.774	1.7
X	7	1:19	15.1	0:11.3	7.065	0.8