VI. Training

A. Introduction

One of the recurrent explanations in the literature for failure to find significant class-size effects is that teachers are teaching the smaller class the same way they have taught a large class and therefore student achievement is not greatly affected.

Most of the research on class size has measured overall effects on student achievement; only a few have examined how teacher behavior changes when class size is reduced (Fox, 1967; Taylor + Fleming, 1972; Wright, et al., 1977, Cahen, et al., 1983; Whittington, et al., 1985.) These studies are summarized in Robinson and Wittebols (1986) as follows: "Research indicates that many teachers, whose classes are reduced in size do not change their teaching techniques (p. 134). Fox found that 55 percent of the teachers made ineffective use of smaller classes; 45 percent did more individualization of instruction. Cahen, et al.(1983) showed that smaller classes had more on-task and engaged time and less time waiting for teacher help. Wright (1977) and Taylor and Fleming (1972) found that teachers in small classes gave students more individual attention. Whittington, et al. (1985) found that teacher logs revealed more individualization, better student behavior, more student participation, and faster pace in smaller classes.

What teachers need to do to teach effectively in a small class does not appear from the literature to differ very much from good teaching in a large class; if they can individualize instruction, increase time on task, and motivate students, the students will learn more. The small class makes these things possible, and a training program should help teachers achieve the possible. Training teachers to work with a teacher aide is a little more complex, because it involves defining roles and developing a teamwork approach to the class.

B. Project STAR Legislation

The Project STAR legislation (Appendix A) specified that teachers should receive in-service education, without specifying the nature or extent of the training. Almost all of the teachers in Project STAR were already involved each year in a variety of in-service education activities. The Project STAR training program would be "in addition to" the in-service training they would normally get. Their usual training varied from school system to school system, and in some systems, individual schools chose what they would do. Teachers with whom we discussed the possibility of a training program agreed that it was a good idea, but it did not seem to be a high priority with most of them. The average teacher in the second grade had 13 years of teaching experience, and the third grade teachers had 14 years of teaching experience. Most of the teachers felt that they knew what and how to teach.

An advisory committee was formed to help plan the training program. The members included training specialists from the State Department of Education, from local systems participating in Project STAR, as well as a superintendent, principal, and teachers from STAR schools.

Vanderbilt University had the major responsibility for planning and conducting the training program statewide, although each of the other universities was involved in the observation of teachers and other data collection activities.

C. Background on Training

1. Training Design

In the winter of 1986-87, 13 of the 75 STAR schools were selected randomly within school types to have their teachers trained. In the selected schools, all of the teachers were trained; there were 57 teachers, 17 percent of the total number in Project STAR. All but two of the initially selected schools agreed to participate, and the two refusals were replaced by other randomly selected schools from the same school type. There were two inner-city schools, two suburban schools, two urban schools, and seven rural schools in the training group. In 1987-88 there were 21 small, 19 regular, and 17 regular/aide classes in the training schools. In 1988-89 there were 25 small classes, 15 regular, and 17 regular/aide classes in the same schools.

In the second grade, 30 percent of the trained teachers had more than 20 years of experience; only 4 of the 57 had less than three years of experience. Within the previous three years about 80 percent had participated in in-service training designed to increase their ability to manage classrooms and increase student learning. As a group, they were not highly motivated to take additional training. The third grade trained teachers were similar in demographic characteristics to the second grade teachers; both groups were similar in demographic characteristics to the Project STAR teachers who were not in the training group.

The training design in both second and third grade involved three days of training before school. Teachers were paid $35 a day to participate. In second grade there were also five, one-hour follow-up sessions, one a month on the average, where the trainers worked with teachers on the improvements they were trying. In the third grade there were three, two-hour follow-up sessions, one each in September, October, and November.

2. Training Curriculum

Dr. Hilda Nason, an experienced trainer, developed the curriculum for the three-day before-school training, with the assistance of an advisory committee of teachers, trainers, and State Department of Education supervisors who reviewed the objectives and made suggestions for the content of the program. The teachers who were to be involved in the training program responded to a questionnaire which listed a number of possible curriculum topics. In addition, the teachers indicated the topics covered by in-service training they had received in the previous three years.

Dr. Nason used these inputs to develop the curriculum and training syllabus which emphasized classroom management, teaching higher order thinking skills, diagnosing students' learning needs, individualizing instruction, and working with aides in the classroom.

The training program was initiated in 1987-88, when the STAR students were in second grade. The second grade teachers in the training schools were observed teaching a reading and a math lesson in the spring of 1987, before they had received any training. Training was provided to five groups of 10 to 15 teachers each in August 1987, before school started. Dr. Nason trained one group of ten teachers. The four other trainers observed Dr. Nason's training sessions. They all used the same training manual that Dr. Nason had developed to try to provide the same content to all teachers. The trainers were all highly rated (over 4.5 on a 1 to 5 scale) in the teacher post-training evaluations of their sessions. Training sessions were observed by other members of the Project STAR staff and the evidence is that training was of uniformly high quality and that the

content of all training sessions was comparable. In the second grade, the same curriculum was delivered to all teachers, regardless of whether they were going to teach a small, regular, or regular/aide class. Teachers had not received their class type assignment at the time they were trained.

The training for the third grade teachers did include one day of condition-specific training and two days of general training in classroom management and teaching thinking skills that was similar to the second grade training. Teachers knew their class-type assignment (small, regular/aide, or regular) at the time of training. Aides participated in the condition-specific day of training, along with their teachers.

A major emphasis in the training in both years was to try to get teacher commitment to implement ideas and concepts from the training in their classrooms. Each teacher was asked to make a specific commitment in writing at the end of the summer training about what they would do in the coming school year to improve their teaching.

In second grade the trainers held live follow-up sessions of about an hour's length each month during the fall at the schools or at a teacher center. In third grade there were three two-hour follow-up sessions in the fall. The trainers also observed each teacher in her or his classroom during the day of one of the follow-up sessions. The follow-up sessions were designed to help teachers implement their commitments to improve their teaching. The follow-up sessions were individualized to respond to teacher interests and did not try to cover a uniform curriculum.

Each teacher in the training program was observed teaching a math lesson and a reading lesson in the fall and again in the winter. For each teacher there was one before-training observation, one observation during the follow-up period in the fall, and one in the winter after the follow-up sessions had been completed. Trainers also did an overall evaluation of the teachers' implementation of the training based on their classroom visits and on their interactions with the teachers during the follow-up sessions.

Observations were also made of 32 teachers in Project STAR schools that were not involved in the training program.

The design permits comparisons between the classroom teaching of trained teachers and of untrained teachers . Comparisons can also be made between observations made before training and after training of the same teachers. Finally, the design permits comparison of teaching behavior of teachers in small classes, regular classes, and classes with a full-time aide. In addition to the observation of teaching, comparisons were made of student achievement between trained teachers and teachers who have not been trained.

3. Observer Training

Observers came to Vanderbilt for a two-day training session prior to spring data collection in 1987. They were given manuals describing the observation system, classroom rating scales, and the data collection instruments and procedures. The observation system provided for descriptive notes as well as for coding specific categories of teacher-student contacts and student task engagement. (See Evertson & Burry, 1989, for a description of the observation system.) Observers practiced using the observation system categories by coding scripted dialogues of teacher-student interactions in class lessons, by contributing their own dialogues for practice, and by coding videotapes of actual class lessons. Throughout training, guidelines for writing "descriptively" rather than "judgmentally" or "evaluatively" were emphasized. At the conclusion of

training, observers used the observation system and the classroom rating scales to record, code, and rate events on a master videotape of a complete class lesson. Criterion-referenced agreement was computed. Agreement with the coded master tape was high (85 percent or above) for the observation system and 80% for the classroom rating scales.

Follow-up contacts with the observers and data from the spring 1987 data collection were used to assess observer agreement. Observers reported little difficulty in using the system since the descriptive notes allowed them to record what they saw and to explain any anomalies that might affect their quantitative data. Prior to the data collection in fall 1987, observers returned for another two-day session. Approximately half of this time was used to talk about classroom events that affected what they recorded. Although agreement was high, observers' scores were not perfect. The primary threat to agreement appeared to be observers' failure to record events because they did not see them, not their failure to interpret observed events accurately. Observer training was conducted again in early September, 1988. At this time, new observers were trained and the skills of those observers who remained with the project for the second year were reinforced.

4. Assessment of Training Effect

To assess the effects of training, trained observers recorded a variety of information (e.g., classroom management, teacher-student interaction, on-task behavior) as teachers taught a reading and a math lesson to their students. Observations were made three times. The first observation was in the spring of the previous academic year. These data provide a pre-training standard. The second observation was in late fall after training (post-training comparison). A third observation was in the winter after a series of at least three follow-up sessions in which the observers met with the teachers to discuss their implementation of the training curriculum.

One primary interest of the project is the effect of training on classroom dynamics. One way to address this issue is to compare the pre-training observations with the post training observations for those teachers who received training. But we concluded that classroom activities at the end of a school year vary considerably from those at the beginning. Hence, the pre-training data do not provide the benchmark for pre- and post- comparisons we would have desired. In other words, any differences in classroom dynamics between pre- and post- data might reflect these naturally occurring differences in the cycle of the school year rather than any real effect of training.

Consequently, we selected eight additional schools whose teachers did not participate in the STAR training program. An effort was made to select schools that were in the same school districts and "school type" (e.g., urban, rural, suburban, or inner city) as those selected for the training program. Selection was constrained somewhat, however, by the proximity of the schools to the trained observers we had available to collect the data. Observers collected the same information on these teachers as on the trained teachers. Hence, the data enable us to assess the effect of training by comparing the observations made of the trained and untrained (or comparison) teachers at the same time in the school year. The comparison group of second grade teachers was observed in the fall of 1987; the third grade group was observed in the fall and winter of 1988-89. Only fall data, then, are available for both academic years for both groups of teachers--trained and comparison.

5. Class Type and Teaching Practices

All of the approximately 340 teachers in both the second and third grades, including the 86 second grade teachers and 84 third grade teachers who were observed for this training study were also randomly assigned to one of three class types: small, regular or regular/aide. Table VI-1 shows the number of teachers by class type for each of the six possible categories of analytical interest for both years of the training study.

Data Analysis

The principal questions of this research are: Does training affect teaching practice (e.g., development of procedures, routines, class management, etc.)?; Do these dynamics vary by class type; and Does any effect of training depend on the particular type of class to which a teacher has been assigned--that is, do class type and training interact? The effect of training on student achievement is also a primary issue. To address these issues, the MEANS and REGRESSION programs available in SPSS-X were employed. MEANS provides the means on variables of interest for each group--trained vs. comparison; small vs. regular vs. regular/aide classes--as well as a test of the significance of the differences. The question of interaction--that is, whether training effects may vary by class type--was addressed with regression analysis using dummy coding.

D. Results

Key variables derived from the category coding section of the observation system are shown in Tables VI-2, VI-3, and VI-4. The observation coding sheet provides for recording teacher-to-student contacts and student-to-teacher contacts in either behavioral, academic, or procedural contexts. Observation time is defined as the actual number of minutes of observation divided by 60. The number of contacts in each category is summed and divided by observation time to obtain a rate per class hour. Proportions of time spent in each activity are calculated by computing the minutes spent in the activity and dividing by observation time. Effect sizes were then calculated for each training group and class type. Variables with effect sizes of .30 or higher are reported. Comparisons among small, regular, and regular/aide classes are shown for math and reading.

All variables from the observation system and classroom rating scale variables are analyzed for each of the three class types and two training conditions for fall observation data only. This time period was selected because the data for the training and comparison groups are the most directly comparable. No data were collected in comparison classes for the winter time period. However, the decision to use these data means that at best we will be attempting to capture the most immediate effects of training on teaching practice in different class types as opposed to the longer term effects that might be captured from using the winter observation data. The winter and fall observation data were compared, however, and very few differences were found for observation variables across the two time periods. Classroom ratings were completed at the end of the observation period for both math and reading lessons; therefore, there are no separate ratings for subject matter for these variables.

The analysis of the classroom rating variables and observation variables was conducted in three steps for each grade. First, the effects of the training program were determined by comparing training and comparison classrooms, regardless of class size. Training effects are summarized in Tables VI-2, VI-3, and VI-4. Next, class type effects were determined by comparing means for small, regular, and regular/aide classes, regardless of training group assignment. Results are

summarized in Tables VI-5, VI-6, and VI-7. Third, interaction effects, that is, the possibility that the effects of training might depend on class type, were also considered. Training and comparison groups were disaggregated by class type in order to test for interaction effects. Significant interaction effects are summarized in Table VI-8. A summary of all effects is presented in Table VI-9.

1. Time Spent in Subject Matter and Lesson Formats (Variables 1-7)

Class means were remarkably similar across class types and grades, both in the amount of time spent in the subject matter and in lesson formats (see Tables VI-6 and VI-7). Class type made no significant difference in the amount of time spent in reading or math; in all class types, more time was spent in reading than in math. Note that the same amount of time spent in a small vs. a regular class can actually result in more time spent per student, meaning that students in small classes may have more turns at the chalkboard in math and more of the teacher's attention in reading.

Reading lesson formats were very similar regardless of class type. In reading lessons, time was spent mainly in small group instruction, with most of the lesson spent in content development, some time spent in independent seatwork, and a much smaller amount of time spent in testing and giving directions for assignments. Third grade classes spent more time in independent seatwork than second grade classes; otherwise, second and third grade reading classes were very similar.

Math lessons across class type and training groups were different from reading lessons but very similar to each other. Small groups were rarely used in math instruction. Most math lesson time was spent in content development, with longer periods of time spent in independent seatwork in math than in reading. These differences in instructional format could contribute to the difference in findings for training and class type effects in reading vs. math.

While no important differences were found among class types, some training effects were apparent in lesson format in math and reading in the second grade (Table VI-3 and VI-4). Trained second grade teachers spent significantly more time in content development in both reading and math than comparison teachers. In reading, trained teachers spent less time in small group instruction than comparison teachers; in math, trained teachers spent less time in independent seatwork than comparison teachers. However, training effects for these variables were not observed in third grade classes.

2. Teacher-Student Contacts (Variables 8-29)

Neither the number nor the types of teacher-student contacts differed by class type in second or third grade reading or in second grade math. Teachers appeared to maintain the same pattern of instruction regardless of class size. In third grade math classes, however, several class type differences were observed (see Tables VI-6 and VI-7).

In third grade math, more of the total contacts were teacher-initiated in regular classes than in small classes. These findings imply that students initiated more contacts in small classes than in regular classes. Means for variable 10 (Table VI-7) support this possibility, although the differences were not statistically significant. The data do indicate that students in small classes may have initiated more procedural contacts than students in regular classes, although this finding was not significant ($p = .10$). Regular classes also had more total academic contacts than regular/aide classes and more questions than either small or regular/aide classes.

The types of contacts teachers initiated varied in some cases by class type. For example, in third grade math, teachers were more likely to initiate questions and make academic contacts in regular classes than in small classes or regular/aide classes.

The same patterns described above for math were observed in third grade reading. However, the differences between means are not statistically significant, with one exception. In third grade reading, small classes and classes with teachers' aides spent a higher percentage of class time in academic activities than did regular classes.

Some training effects were apparent in teacher-student contact variables in reading classes, but not in math classes (Table VI-3 and VI-4). More of the total contacts in training classes were questions, and fewer of the total contacts in second grade training classes were directives (direct statements to students that require them to respond as opposed to questions or comments). Trained third grade teachers initiated fewer questions than comparison teachers. Trained second grade teachers initiated more comments than comparison teachers. Fewer of the total contacts in second grade training classes were directives.

In addition to the direct effects of class type and training noted, some interaction effects also surfaced. Training effects varied depending on class type on two variables (see Table VI-8). In second grade math, trained teachers in small and regular classes had fewer procedural contacts than comparison teachers, but trained teachers with classroom aides had more procedural contacts than comparison teachers. In third grade reading, trained teachers in small classes and classes with teacher aides made more directive contacts than comparison teachers; in regular classes, however, trained teachers made fewer directive contacts than comparison teachers. These apparent interactions are important to consider because they can explain why direct effects of either training or class type are sometimes nonsignificant.

3. Student Outcome Variables (Variables 30-33)

Training and class type each had effects on student task engagement. Direct class type effects were present in second grade reading, where students in small and regular/aide classes had fewer students probably on-task and more students definitely on task than students in regular classes (see Tables VI-6 and VI-7). A similar effect was observed in third grade reading, where significantly more students were probably on-task in training classes than in comparison classes, with no significant difference in the percentage of students definitely on-task. These results were probably related in part to the finding in grade 2 that small classes have better visibility than regular classes: it may have been easier for observers to see whether students were definitely on-task in the less-crowded classrooms.

As Table VI-8 shows, the two variables training and class size interacted on the percent of students probably on-task (Variable 31), with the highest percentage of students probably on-task in regular training classes. Variation in on-task behavior due to training depends on the class type observed. Thus, unique combinations of training condition and class type contributed to the observed effect in the second grade.

An interaction between training and class size may have masked an effect on the percent of students waiting in second grade reading. Fewer students were observed waiting in small classes and classes with teacher aides, although this difference was not significant ($p = .10$). Table VI-8 shows that, while there was no difference on this variable in small training or comparison classes, training had opposite effects in regular classes than in classes with teacher aides. Thus it appears that unique combinations of training and class size affected the percent of students waiting.

4. Classroom Rating Scales (Variables 1-31)

Training appears to have been more important than class type on the classroom rating scale variables, although some results of training are conflicting. There were no class type effects for classroom rating variables in grade 3; in grade 2, small classes received higher ratings for suitable traffic patterns and for greater visibility (Table VI-5).

Training, however, had several effects on classroom rating variables (Table VI-2). In the second grade, trained teachers' classrooms functioned more smoothly. Trained teachers organized their classrooms for better visibility, used efficient routines, procedures, and transitions, and had needed materials ready. These teachers described their objectives more clearly than comparison teachers. Surprisingly, however, their students exhibited more avoidance behavior during seatwork. All significant training effects in the third grade were negative, an unexpected result. These teachers were rated lower for their pacing of lessons, had a less task-oriented focus, and gave explanations and presentations that were less clear than comparison teachers. Training for third grade teachers was condition-specific and training for second grade was general.

5. Findings from Observer Narrative Descriptions

As part of the observation protocol, observers kept narrative records of the classes they visited. The records provide useful contextual information about the ratings given and the effects observed. A preliminary review of all third-grade narratives and a random selection of second-grade narratives yields a wealth of information about the observed classes.

The lack of consistent results between second and third grades prompts the question of whether classroom processes and curricula between the two years are qualitatively different. For the most part, the classes appear to be fairly similar, especially in reading. The format of teacher-led small - group instructions in reading predominates in both grades and all class types. In math some differences are evident. Second grade training teachers used more manipulatives of various types in presenting content than did second grade comparison teachers or third grade teachers. Second grade classes spent more time working through problems and worksheets as a class and less time in independent seatwork than did third grade classes. One reason for this difference may have been teachers' attempts to accommodate the shorter attention spans of second graders; observers commented much more frequently on wiggling, chair-twisting, and other physical expressions of excess energy in the second-grade narratives.

More marked than the differences between second and third grades were the similarities between classes, both between and within grades. A reader given an unmarked narrative would be hard pressed to decide whether it was a regular or small class. Teachers apparently made few changes in curriculum, lesson format, or methods based on class size. It might be expected that smaller classes would have smaller reading groups, but this does not appear to have been the case. Teachers in larger classes often had at least one reading group of three or four (this was frequently the lowest reading group), and often met with three or four groups rather than with the one or two groups common in the small classes. In some small classes, teachers took advantage of the lower numbers by meeting with the class as one large reading group. Also, most larger classes had several students who left during reading, presumably for remediation, reducing so-called "regular" classes to the size of the small classes. This was especially common in the third grade. The presence of an aide made little difference in the numbers in groups, as the aide usually either monitored seatwork or accompanied students to resource classes rather than working with a separate reading group.

Within grades, math classes were also highly similar. Teachers presented or reviewed a concept by having students work problems at the board or at their desks; the class worked through a worksheet together, and problems were assigned to be completed individually (usually while reading groups met). Math lessons ended when it was time for lunch. Working in groups was extremely rare, regardless of class size; use of manipulatives was infrequent, especially in the third grade.

Teachers used aides in two distinct ways. About half of the aides did little other than clerical work. They monitored the class if the teacher was called out of the room, but otherwise they had only limited contact with students. Other aides were much more active with students, circulating while the teacher presented math content and during seatwork assignments and occasionally meeting with a reading group or with individual students during reading. One exception to the limited use of aides with reading groups occurred in one training classroom. Students in this classroom were divided into three reading groups. At any given time, one group worked with the aide, one group worked with the teacher, and one group worked on seatwork. Each group met with both the aide and the teacher. Thus, each student had one hour of direct reading instruction rather than the usual thirty minutes. This was the only observed classroom where the aide was used so fully.

Content and format of the third grade lessons was surprisingly consistent across classrooms. Teachers clearly kept closely to the prescribed curriculum. The shortcomings of the curriculum were clear. The focus in almost all classes was on task completion rather than on understanding concepts. Lessons consisted of unrelated pieces of information rather than units of meaning. This was especially evident in reading, where a typical lesson might include vocabulary drill, dictionary skills, phonics, and oral reading. Teachers occasionally related vocabulary words to reading; only very infrequently did teachers relate the phonics they had just been drilling to the stories students read. Very few teachers had students write anything more than the few words necessary to complete a worksheet. The isolation of the teaching of reading skills from reading itself is typified by the casual comment of one teacher to the observer that her class wasn't having reading today--they were going to the library instead!

A review of the narratives from math and reading classes reveals classes that are remarkably similar, regardless of training or class type. Teachers rarely waver from the curriculum they are given, which stresses skills in isolation from meaning.

6. Training and Student Achievement

In both second and third grades, classes with trained teachers had slightly higher scores in both reading and math than the classes with untrained teachers (see Table VI-II). In second grade, trained teachers in each class type had higher scores than untrained teachers. In third grade, the untrained teachers in small classes had higher mean scores than the trained teachers, but the trained teachers had higher class averages in the other two class types. While some of the second grade differences approached significance, overall training did not make a significant difference in student achievement in either second or third grade.

An important comparison is of differential growth in achievement of students in classes with trained teachers as compared to untrained teachers, because this adjusts for differences that may exist in the beginning test scores. In the second grade, trained and untrained teachers had very similar gain scores in both reading and math, and in the third grade, untrained teachers had slightly higher gain scores in both subjects, although the differences were not significant (see Table VI-12). Gains by class type and training exhibited inconsistent patterns between reading

and math and between the second and third grades. The differences in the gain scores of trained and untrained teachers were small and nonsignificant. While there were some interactions between class type and training, they were inconsistent across subjects (reading and math) and second and third grades.

The overall conclusion is that training does not make a significant difference in student achievement, nor does it make a significant difference for any one class type. Training did not help the small class teachers to improve student achievement any more than it helped regular class teachers or teachers with an aide.

One hypothesis is that some teachers respond to training and make changes in their teaching styles or try new things, while others do not. The trainers rated the teachers at the end of their three-day training on a five-point scale on their attitude toward the training, their participation in the training, and their commitment to try some new things. Trainers gave teachers very positive ratings for attitude, positive for participation, but fairly neutral for commitment. These teacher ratings were correlated with student gains in reading and math. There was a positive correlation of about .4 between the trainers' ratings of the teachers' attitudes, participation and commitment and student achievement in reading. For math, the correlation was lower, .17 for attitude, .01 for participation, and .16 for commitment (second grade data). These positive correlations between attitude and commitment estimates and subsequent class performance suggest that teachers with good attitudes about teaching and who are willing to make commitments to try new things are likely to be effective teachers. Training, however, doesn't necessarily "cause" student achievement; and it may be that these qualities existed before the training.

Trainers also rated the teachers at the end of the follow-up period on the extent to which they had responded to the training. About 75 percent were rated as responsive, while 25 percent did not respond to training. This was a subjective rating, and some of the trainers gave their teachers more positive ratings than others, so there is a definite trainer effect on the ratings. The mean achievement level in classes where the teacher responded to the training was not significantly different from the mean class achievement level of the teachers who did not respond to the training. There was also no significant difference in their gain scores.

7. Discussion of the Observation Result

There are several possible reasons why the training provided to the second and third grade teachers did not lead to significant improvement in student average test scores or gains. The great majority of Project STAR teachers were experienced and about four-fifths of them had participated in other in-service training within the preceding three years. It is clearly not accurate to refer to the remainder of the STAR teachers as "untrained, " for most of them had some similar training. Therefore, in an experienced group of teachers the marginal effect of three days of additional training may not be large enough to affect student test scores.

Second, the training emphasized topics such as teaching higher order thinking skills and diagnosing students' learning needs. Even if the teachers benefitted from the training and were able to apply the skills that they were taught, this might not be reflected in the test results. Tests focused on "basics."

Third, many of the teachers were not highly motivated to participate in the training. In the exit interviews when the teachers were asked if the training program led them to change their teaching in any way, about half the teachers said that it did not. Of those who did say it helped them, several gave general answers such s "it made me more creative," "I tried some new things

and they helped some. " A few made specific comments about how they had taught higher order thinking skills. When asked whether they thought the changes they had made in teaching would be reflected in their students' test scores, less than 10 percent said yes.

The interviews at the end of the year were in marked contrast to the very positive teacher ratings of the training at the time they completed it and the initial commitment that most of the teachers made to try something different. This suggests that training is not sufficiently reinforced, even when there are follow-up sessions with a skilled trainer, to get a majority of teachers to incorporate it into their classroom repertoire of skills and procedures. It may not be possible for even a highly skilled outside trainer to encourage experienced teachers to do something new unless the teachers are self-motivated to improve. In addition, the improvement effort needs to be strongly reinforced by the principal and/or local system supervisors. While a number of the principals had a positive attitude about the training program, most of them were uninvolved. To try to stimulate more school and system reinforcement for the training, Hilda Nason, who had developed the training package, visited all of the training school principals in the summer before the third- grade training. She discussed the objectives and methods of the training program with the principals and tried to involve them in reinforcing their teachers' improvement efforts. The evidence from the third grade results suggests that this strategy did not make a measurable difference in student achievement.

One limitation to the study of which the investigators were aware from the outset was the relatively small amount of observation time available. However, even with the limited observation time, there appears to be support for the effects of class size on teacher and student behavior in grade 2 but not grade 3. There are predictable differences in class processes that follow simply from the numbers: students are more visible; each student is more likely to get a turn more often during class lessons; students do not have to wait as long for help; students can initiate more contacts with teachers.

The unique feature of this study was the inclusion of training for a subsample of teachers and there are training effects particularly for practices related to classroom management (e.g., the efficiency of classroom routines, general procedures, transitions). Also, teachers in all types of classes appear to benefit equally from training, although this effect was not strong.

Statistical findings for differences in teacher behavior between class sizes and for trained and untrained teachers were not strong and formed no coherent pattern of effects. Several significant findings in grade 2 were contradicted in grade 3: variables with significant effects in grade 2 were not significant in grade 3 or, in some cases, actually showed opposite effects in grade 3.

The most important findings had to do with similarities rather than differences. Teachers of both grades and all class types spent much more time in reading than in math. Descriptive notes provide insights into how instruction occurs in these highly similar lessons. Teachers orchestrate a narrow, tightly controlled skills approach to the curriculum. There is very little variation from this model, indicating that the State Basic Skills First curriculum had the effect of making lesson content and format more uniform. While this study controlled for training, class size, follow-up, and feed back to teachers, it did not control for the nature of the curriculum, and it is clear from narrative descriptions that curriculum, especially in reading, exercised strong influence on the way teachers taught. The same lessons appeared repeatedly across classes and cross schools in both reading and math. Clearly these teachers were following the curriculum so closely that they were virtually in the same place in the book at the same time. This finding leads to a series of questions regarding what possible effects training or class size can be expected to have when the requirements of the curriculum clearly dominate the pacing and structure of classroom lessons.

Our data point up the importance of the curriculum to the type of learning we are encouraging in our schools. If the content we desire is a basal-driven, isolation of skills and memorization of facts as the core for second and third grade reading and math, then the strategies teachers have developed to cope with their curriculum are satisfactory. However, if instructional goals are to increase the development of higher-order thinking skills, creativity, and personal responsibility for learning, a reduced teacher/student ratio may be important to enable teachers to achieve these objectives effectively. Fewer rote tasks, more reading and writing in context, more problem-solving activities -- all will require more teacher/student interaction than the present curriculum. If such broad changes in learning outcomes are desired, changing class size and training teachers alone will not be enough; these changes must be coupled with a curriculum focused on these objectives.

Table VI-1

Teachers in Training, Comparison and Other Groups in Grades 2 and 3 by Class Types, Project STAR

Second Grade	Trained	Comparison	Others	Total
Small	23	13	97	133
Regular	14	10	76	100
Regular/Aide	17	9	81	107
Total	**54**	**32**	**254**	**340**

Third Grade	Trained	Comparison	Others	Total
Small	25	15	100	37
Regular	15	8	67	21
Regular/Aide	17	9	80	26
Total	**57**	**32**	**257**	**336**

Figure VI-1

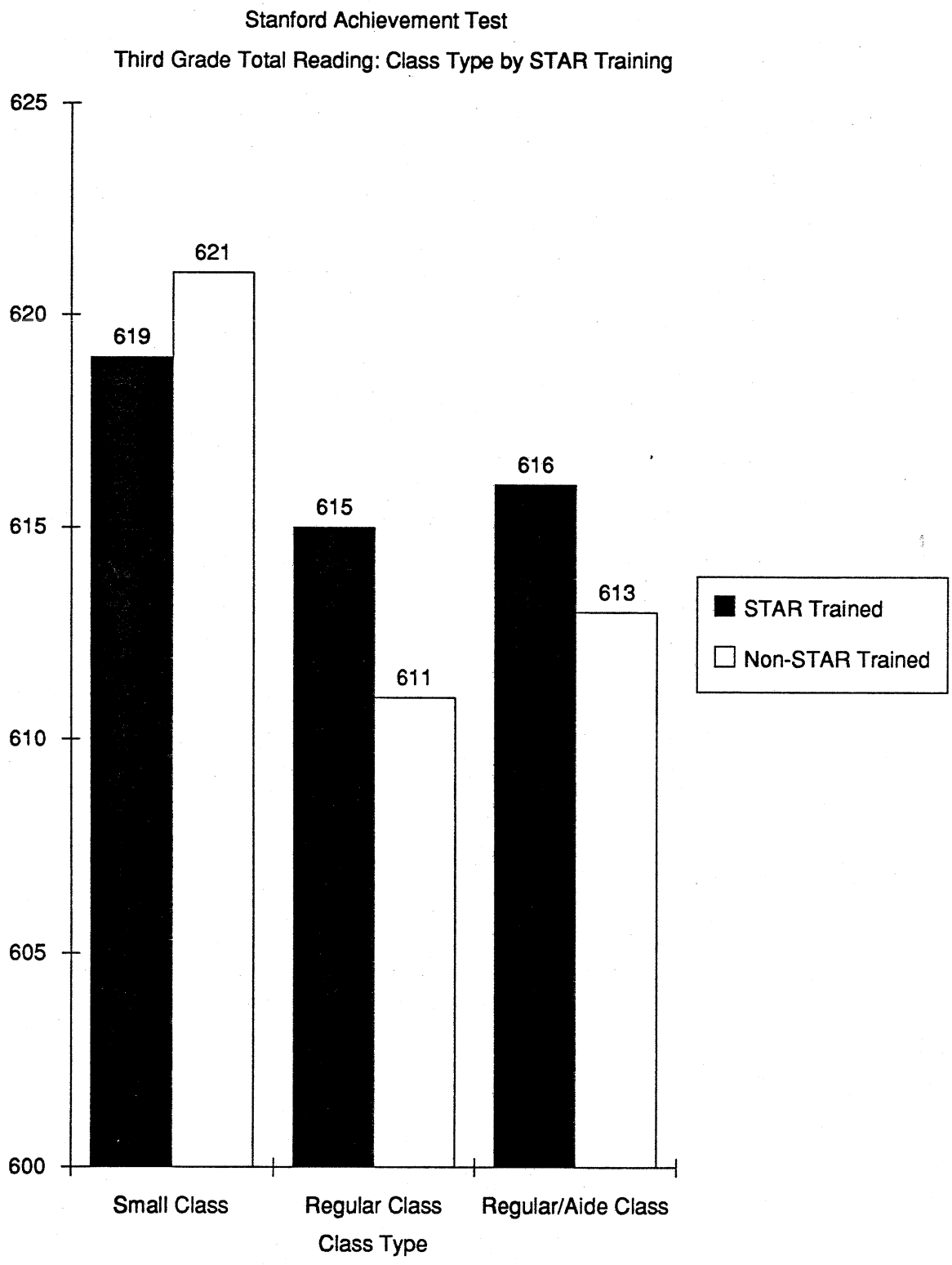
Stanford Achievement Test
Third Grade Total Reading: Class Type by STAR Training
625
620
615
610
605
600
619
621
615
611
616
613
STAR Trained
Non-STAR Trained
Small Class
Regular Class
Regular/Aide Class
Class Type

TABLE VI-2. Means and Standard Deviations for Grade 2 and Grade 3 Training Conditions Aggregated Across Class Type. Classroom Rating Variables

VARIABLE	GRADE 2-FALL ONLY						GRADE 3 - FALL ONLY					
	TRAINING (n=54)		COMPARISON (n=32)				TRAINING (n=53)		COMPARISON (n=31)			
(Classroom Rating)	M	SD	M	SD	Sig.	Effect Size	M	SD	M	SD	Sig.	Effect Size
1. Suitable Traffic Patterns	4.57	.72	4.25	.76	.06	.43	3.92	.86	4.16	.78		
2. Good Visibility	4.54	.69	4.22	.79	.05	.40	3.79	.97	4.13	.76	.10	-.44
3. Describes Objectives Clearly	4.63	.71	4.22	.79	.01	.52	3.81	.93	4.03	.91		
4. Materials are Ready	4.69	.58	4.34	.75	.02	.46	4.04 (n=52)	.83	4.32	.83	.13	-.34
5. Clear Directions for Assignments	4.52	.77	4.50	.67			3.94	.86	4.23	.88	.16	-.32
6. Individualized Assignments	3.67 (n=46)	1.03	3.56 (n=25)	.87			2.89 (n=38)	1.29	2.96 (n=26)	1.15		
7. Provides, Seeks Rationales	4.11	.85	4.03	.65			3.40	1.12	3.77	1.15	.14	-.33
8. Appropriate Pacing of Lesson	4.52	.77	4.50	.76			3.68	1.05	4.13	.96	.05	-.47
9. Clear Explanations, Presentations	4.57	.77	4.48 (n=31)	.63			3.85	.93	4.29	.86	.03	-.51
10. Monitors Student Understanding	4.50	.82	4.34	.75			3.81	.96	4.13	.92		-.34
11. Enforces Work Standards	4.63	.65	4.47	.80			3.79	1.03	3.97	1.02		
12. Efficient Admin. Routines	4.55 (n=47)	.69	4.19 (n=27)	.88	.05	.42	3.85	.99	4.10	1.11		
13. Appropriate General Procedures	4.54	.82	4.13	.79	.02	.52	3.85	.89	4.10	.94		
14. Efficient Small Group Procedures	4.56	.79	4.41	.71			3.86 (n=43)	1.10	4.26 (n=27)	1.02	.13	-.39
15. Routines for Academic Work	4.53 (n=51)	.73	4.30 (n=30)	.75		.31	3.87	1.00	4.03	1.05		
16. Considers Attention Spans	4.24	.78	4.13 (n=30)	.94			3.60 (n=52)	.96	3.94	.85	.11	-.40
17. Successful Students	4.43 (n=53)	.67	4.30 (n=30)	.60			3.89	.85	4.16	.82	.15	-.33
18. Actions Related to Students' Interest	4.37	.65	4.10 (n=30)	.76	.09	.36	3.57	.95	3.90 (n=30)	.92	.12	-.36

19. Rewards Good Performance	4.46	.88	4.42	.99			3.74	1.02	4.10	1.08	.13	-.34
			n=31									
20. Consistent	4.47	.87	4.26	.86			3.94	.86	4.06	1.03		
	n=53		n=31									
21. Effective Monitoring	4.46	.86	4.26	.97			3.83	.99	3.90	1.19		
			n=31									
22. Efficient Transitions	4.48	.75	4.06	1.06	.04	.39	3.71	1.00	4.13	.99	.07	-.42
			n=31				n=52					
23. Disruptive Behavior	1.72	.91	1.71	1.05			1.75	.89	1.59	.89		
	n=53		n=28				n=48		n=27			
24. Stopped Quickly	4.24	1.15	4.48	1.12			4.00	1.13	4.35	1.09		
	n=33		n=21				n=42		n=20			
25. Ignored	2.69	1.51	1.93	1.38	.12	.55	1.79	1.08	2.45	1.29	.11	-.51
	n=29		n=14				n=29		n=11			
26. Inappropriate Behavior	1.75	.83	1.77	.96			1.82	.73	2.00	1.13		
	n=53		n=31				n=49		n=26			
27. Stopped Quickly	4.38	1.01	3.88	1.48	.12	.34	3.98	1.16	4.33	1.01		
	n=37		n=25				n=46		n=24			
28. Ignored	2.79	1.45	2.00	1.20	.07		2.06	1.22	2.36	1.29		
	n=33		n=15				n=33		n=11			
29. Task-Oriented Focus	4.70	.57	4.45	.72	.08	.35	3.77	1.03	4.32	.87	.01	-.63
			n=31									
30. Relaxed, Pleasant Atmosphere	4.44	.74	4.45	.89			3.81	1.14	4.29	1.01	.06	
			n=31									
31. Avoidance Behavior During Seatwork	2.70	1.53	2.06	1.18	.05	.54	2.61	1.31	2.93	1.47		
	n=53		n=31				n=51				n=27	

n's are indicated when missing cases > 10%.

TABLE VI-3. Means and Standard Deviations for Grade 2 and Grade 3 Training Conditions Aggregated Across Class Type. Reading Variables.

VARIABLE	GRADE 2-FALL ONLY						GRADE 3 - FALL ONLY					
	TRAINING (n=54)		COMPARISON (n=32)				TRAINING (n=53)		COMPARISON (n=31)			
(READING)	M	SD	M	SD	Sig.	Effect Size	M	SD	M	SD	Sig.	Effect Size
1. Av. Time in Subject (min.)	73.38	24.99	75.51	26.84			66.48	25.06	66.52	27.21		
2. No. of Small Groups	2.70	1.42	2.69	1.42			1.44	1.23	1.42	1.15		
3. Content Development	14.21	24.55	1.95	3.43	.01	3.57	10.71	15.10	13.58	17.81		
4. Assignment Directions	1.14	2.62	1.69	4.48			1.30	2.46	1.72	2.83		
5. Indep. Seatwork	1.65	3.94	2.01	4.02			10.30	15.46	7.21	12.31		
6. Small Group Instruction	36.77	25.73	50.39	15.03	.01	-.91	31.47	24.70	31.63	25.72		
7. Testing	1.04	7.68	1.15	6.49			.45	1.87	.58	2.25		
8. % Indiv. Contacts	69.96	27.97	67.54	35.94			68.74	34.30	75.87	35.06		
9. % Teacher Init. Contacts	62.21	24.87	60.14	34.48			53.19	24.00	61.31	30.04		
10. % Student Init. Contacts	8.31	7.75	7.36	8.03			15.55	18.76	14.56	13.47		
11. % T-Initiated Directives	19.23	13.93	27.55	30.74	.09	15.93	15.15	13.88	11.23			
12. % T-Initiated Questions	36.97	19.89	29.90	19.45	.11	.36	31.85	18.25	42.56	24.91	.03	-.43
13. % T-Initiated Comments	5.78	6.11	2.62	3.41	.01	.93	5.40	6.71	4.87	8.64		
14. % T-Initiated Ac. Contacts	49.93	22.96	51.80	31.31			42.82	22.03	49.97	27.86		
15. % T-Initiated Behav. Contacts	5.53	4.43	4.93	6.07			4.06	4.43	4.81	7.40		
16. % S-Initiated Questions	4.41	4.32	4.16	5.95			9.59	12.68	6.95	8.36		
17. % S-Initiated Comments	3.90	4.62	3.45	4.24			5.96	8.55	7.61	8.35		
18. % S-Initiated Ac. Contacts	5.75	5.66	5.78	6.42			10.02	12.18	10.62	8.99		
19. % S-Initiated Proc. Contacts	2.59	3.24	1.66	2.47	.17	.38	5.33	9.27	3.94	5.80		
20. #/hr. Contacts Praise *	1.23	2.32	.58	1.82	.18	.35	..	..	..	..		
21. #/hr. Contacts Criticism *	2.72	3.17	1.52	2.35	.07	.51	..	..	..	..		

VARIABLE	TRAINING (n=54)		COMPARISON (n=32)				TRAINING (n=53)		COMPARISON (n=31)			
(READING)	M	SD	M	SD	Sig.	Effect Size	M	SD	M	SD	Sig.	Effect Size
22. % Contacts T-Initiated	89.70	8.09	86.73	17.42			79.53	19.81	80.81	14.27		
23. % Contacts Academic	79.54	14.66	85.27	10.65	.06		78.64	18.84	80.90	16.57		
24. % Contacts Behav.	7.24	6.68	6.36	5.59			5.66	6.00	6.70	9.80		
25. % Contacts Directives	27.52	16.11	39.01	19.09	.004		23.12	18.80	18.96	14.78		
26. % Contacts Questions	58.96	19.08	50.35	20.92	.05		61.99	20.35	65.72	17.45		
27. % Contacts Praise *	1.88	3.52	.90	2.66	.18	.37	..	..	..	..		
28. % Contacts Criticism *	4.12	4.92	2.31	3.12	.07	.58	..	..	..	..		
29. % Time Spent in Academic Acts.	97.64	3.69	97.43	3.36			93.61	9.03	93.68	6.69		
30. % Students Def. On-Task	86.72	9.12	89.52	9.36			86.38	12.12	88.48	13.48		
31. % Students Prob. On-Task	4.30	3.32	3.22	2.53	.11	.43	5.91	7.54	3.39	4.69	.10	.56
32. % Students Off-Task	6.90	6.47	5.52	6.55			5.97	5.78	5.61	8.20		
33. % Students Waiting	1.78	2.92	1.70	2.72			2.19	4.55	2.05	4.04		

* Data for this variable not collected for Grade 3.

VI-4. Means and Standard Deviations for Grade 2 and Grade 3 Training Conditions Aggregated Across Class Type. Math Variables.

VARIABLE	GRADE 2 · FALL ONLY						GRADE 3 · FALL ONLY					
	TRAINING (n=54)		COMPARISON (n=32)				TRAINING (n=53)		COMPARISON (n=31)			
(MAIN)	M	SD	M	SD	Sig.	Effect Size	M	SD	M	SD	Sig.	Effect Size
1. Av. Time in Subject (min.)	38.42	10.60	34.20	14.37			40.04	15.35	40.93	15.24		
2. No. of small groups	.30	1.33	.06	.35			.27	.53	.17	.38		
3. Content Development	37.86	17.05	29.86	20.05	.04	.40	32.24	16.64	30.95	14.53		
4. Assignment Directions	3.35	7.52	4.50	6.67			4.36	6.55	2.76	3.17		
5. Independent Seatwork	11.98	12.65	18.04	17.08	.06	-.36	14.97	15.20	17.20	13.77		
6. Small Group Inst.	.36	2.67	.52	2.07			.41	2.97	.54	2.95		
7. Testing	3.26	11.76	3.74	11.49			.96	3.29	.22	1.22		
8. % Indiv. Contacts	71.93	41.86	66.55	43.72			67.67	39.71	64.20	40.06		
9. % T-Initiated Contacts	63.25	38.33	55.48	39.32			55.64	31.43	52.10	30.24		
10. % S-Initiated Contacts	8.52	9.18	9.80	11.99			12.04	17.25	12.10	14.91		
11. % T-Initiated Directives	17.80	15.35	20.62	25.29			17.12	20.99	14.72	14.72		
12. % T-Initiated Questions	34.67	31.53	27.56	26.39			30.24	19.76	30.96	19.91		
13. % T-Initiated Comments	8.74	9.63	7.30	8.10			8.28	11.48	6.42	12.06		
14. % T-Initiated Ac. Contacts	46.90	32.24	44.98	34.05			42.82	25.16	40.83	23.63		
15. % T-Initiated Behav. Contacts	6.74	5.99	5.60	9.55			3.82	6.57	5.42	7.87		
16. % S-Initiated Ques.	5.99	7.06	6.29	8.74			9.21	15.03	8.54	10.76		
17. % S-Initiated Comments	2.68	3.44	3.18	6.02			2.82	3.56	4.22	6.42		
18. T S-Initiated Ac. Contacts	6.07	7.44	6.55	8.15			7.75	11.75	7.68	9.66		
19. % S-Initiated Proc. Contacts	2.55	3.27	3.24	5.52			4.29	7.81	4.41	6.62		
20. #/hr. Contacts Praise *	2.30	4.44	1.58	6.56			..	..	..	..		
21. #/hr. Contacts Criticism *	2.23	3.06	1.40	2.12			..	..	..	..		

VARIABLE	TRAINING (n=54)		COMPARISON (n=32)				TRAINING (n=53)		COMPARISON (n=31)			
(MATH)	M	SD	M	SD	Sig.	Effect Size	M	SD	M	SD	Sig.	Effect Size
22. % Contacts T-Initiated	82.21	24.02	80.20	22.20			84.43	18.23	85.32	14.92		
23. % Contacts Academic	70.90	23.64	79.04	17.90	.10	<u>-.45</u>	76.29	20.96	80.58	18.25		
24. % Contacts Behavioral	9.13	7.69	7.70	10.05			5.67	8.19	7.23	9.60		
25. % Contacts Directives	27.99	20.50	29.41	23.10			23.30	20.91	28.04	24.53		
26. % Contacts Questions	55.55	22.37	54.92	25.96			60.29	25.90	61.46	24.66		
27. % Contacts Praise *	3.84	6.84	2.91	9.74			--	--	--	--		
28. % Contacts Criticism *	3.81	5.83	2.83	4.50			--	--	--	--		
29. % Time Spent in Academic Acts.	98.53	4.06	96.50	8.31			<u>95.01</u>	<u>9.02</u>	<u>94.06</u>	<u>11.20</u>		
30. % Students Def. On-Task	87.25	9.99	87.04	12.37			85.22	15.44	88.31	15.57		
31. % Students Prob. On-Task	3.86	2.99	3.12	4.63			4.38	6.23	1.71	3.14	<u>.03</u>	<u>.85</u>
32. % Students Off-Task	4.80	7.05	4.55	7.50			6.62	9.19	5.55	9.59		
33. % Students Waiting	4.43	6.52	5.83	8.46			3.66	7.84	4.44	10.09		

* Data for this variable not collected for Grade 3.

TABLE VI-5. Means and Standard Deviations for Grade 2 and Grade 3 Class Types Aggregated Across Training Conditions. Classroom Rating Variables.

VARIABLE (Classroom Rating)	GRADE 2 SMALL (n=36) M	SD	REGULAR (n=24) M	SD	AIDE (n=26) M	SD	SIG.	GRADE 3 SMALL (n=37) M	SD	REGULAR (n=21) M	SD	AIDE (n=26) M	SD	SIG.
1. Suitable Traffic Patterns	4.69	.58	4.38	.71	4.19	.94	.03	4.11	.66	4.19	.93	3.73	.92	
2. Good Visibility	4.69	.58	4.21	.78	4.23	.82	.01	4.05	.85	4.00	.84	3.65	1.02	
3. Describes Obj. Clearly	4.47	.74	4.54	.88	4.42	.70		4.03	.90	3.95	.94	3.65	.94	
4. Materials Are Ready	4.53	.70	4.67	.56	4.50	.71		4.24	.76	4.10	.89	4.04	.92	
5. Clear Dir. for Assignments	4.44	.77	4.54	.72	4.58	.70		4.11	.84	4.19	.75	3.85	1.01	
6. Individualized Assignments	3.55	.91	3.90	.94	3.48	1.08		2.80	1.22	3.05	1.32	2.95	1.18	
7. Provides, Seeks Rationales	4.11 (n=29)	.82	4.22 (n=21)	.52	3.92 (n=21)	.89		3.76 (n=25)	1.26	3.45 (n=20)	.98	3.27 (n=19)	1.04	
8. Appropriate Pacing of Lesson	4.53	.84	4.46	.66	4.54	.76		3.92	.95	3.95	1.12	3.65	1.09	
9. Clear Exp., Presentations	4.61	.69	4.54	.72	4.44	.77		4.14	.89	4.10	.83	3.77	1.03	
10. Monitors Students Und.	4.50	.81	4.25	.94	4.53	.58		4.00	.94	4.10	.77	3.69	1.09	
11. Enforces Work Standards	4.64	.76	4.38	.77	4.65	.56		3.95	.94	4.05	1.02	3.58	1.10	
12. Efficient Admin. Routine	4.34	.81	4.45	.74	4.48	.79		4.08	.95	4.00	1.00	3.69	1.16	
13. Appropriate General Proc.	4.39 (n=29)	.90	4.42 (n=22)	.72	4.35 (n=23)	.85		4.08	.76	4.00	.84	3.69	1.12	
14. Efficient Small Group Proc.	4.50	.91	4.50	.66	4.50	.65		4.07	.98	4.17	1.04	3.82	1.26	
15. Routine for Academic Work	4.42	.83	4.39	.66	4.52	.71		3.97 (n=30)	1.01	4.00 (n=18)	.89	3.81 (n=22)	1.13	
16. Considers Attention Span	4.34	.84	4.09	.73	4.12	.91		3.81	.97	3.86	.91	3.48	.87	
17. Successful Students	4.40	.65	4.35	.65	4.40	.65		4.11	.77	4.05	.86	3.77	.91	
18. Actions Rel. to Studs.' Int.	4.23	.77	4.35	.65	4.27	.67		3.78	.98	3.75	.97	3.50	.91	
19. Rewards Good Performance	4.31	1.17	4.43	.79	4.65	.56		3.97	.96	3.81	1.08	3.77	1.18	
20. Consistent	4.47	.81	4.09	1.11	4.54	.65		4.08	.83	4.05	.92	3.81	1.06	
21. Effective Monitoring	4.39	.90	4.13	1.14	4.62	.57		3.95	.97	3.86	1.01	3.73	1.25	
22. Efficient Transitions	4.31	.98	4.43	.84	4.27	.83		3.92	.86	4.10	.94	3.60	1.22	

VARIABLE (Classroom Rating)	GRADE 2 SMALL (n=36) M	SD	REGULAR (n=24) M	SD	AIDE (n=26) M	SD	SIG.	GRADE 3 SMALL (n=37) M	SD	REGULAR (n=21) M	SD	AIDE (n=26) M	SD	SIG.
23. Disruptive Behavior	1.66	1.00	1.95	1.13	1.58	.65		1.73	1.01	1.78	.94	1.58	.65	
24. Stopped Quickly	4.59	.80	4.13	1.26	4.19	1.38		4.24	1.02	3.73	1.49	4.22	.88	
25. Ignored	2.17	1.38	3.00	1.53	2.25	1.60		1.76	1.25	2.50	1.35	1.85	.80	
26. Inappropriate Behavior	1.58	.77	2.04	1.11	1.76	.72		1.67	.82	1.94	1.00	2.13	.85	
27. Stopped Quickly	4.46	1.06	3.95	1.31	4.05	1.35		4.19	1.14	4.00	1.28	4.05	.97	
28. Ignored	2.20	1.28	3.07	1.44	2.50	1.51		2.00	1.33	2.40	1.51	2.13	.96	
29. Task-Oriented Focus	4.56	.73	4.61	.66	4.69	.47		4.05	1.00	4.05	1.02	3.81	1.02	
30. Relaxed Pleasant Atmos.	4.44	.88	4.52	.67	4.38	.80		4.03	1.09	4.19	1.17	3.76	1.11	
31. Avoidance Behav. Dur. Seat.	2.31	1.45	2.61	1.44	2.56	1.45		2.60	1.52	2.72	1.32	2.88	1.20	

n's are indicated when missing cases > 10%.

TABLE VI-6. Means and Standard Deviations for Grade 2 and Grade 3 Class Types Aggregated Across Training Conditions. Reading Variables

VARIABLE (READING)	GRADE 2 SMALL (n=36) M	SD	REGULAR (n=24) M	SD	AIDE (n=26) M	SD	SIG.	GRADE 3 SMALL (n=37) M	SD	REGULAR (n=21) M	SD	AIDE (n=26) M	SD	SIG.
1. Avg. Time in Subject	70.63	22.00	77.08	27.22	76.38	28.82		64.22	24.59	72.47	28.42	65.27	25.59	
2. No. of Small Groups	2.69	1.69	2.54	.83	2.85	1.46		1.56	1.18	1.21	1.18	1.42	1.24	
3. Content Development	10.73	21.10	10.15	21.32	7.68	19.12		11.47	16.26	13.74	14.40	10.87	17.61	
4. Assignment Directions	.72	1.90	1.63	4.14	1.94	4.23		1.18	2.29	1.64	2.51	1.72	3.08	
5. Independent Seatwork	2.57	5.32	1.34	2.97	1.10	1.92		11.46	17.77	9.10	10.66	5.89	10.72	
6. Small Group Instruction	40.53	23.09	41.33	25.17	44.11	22.25		32.18	25.90	25.34	24.22	35.16	24.18	
7. Testing	.00	.00	.00	.00	3.58	12.96		.51	2.13	.74	2.30	.32	1.65	
8. % Individual Contacts	66.58	30.42	68.38	29.06	73.11	34.14		73.24	35.58	72.24	38.69	68.44	30.92	
9. % T-Init. Contacts	57.66	26.84	61.43	25.77	66.67	33.48		56.96	26.95	59.50	30.57	53.03	23.54	
10. % S-Init. Contacts	8.94	8.66	8.19	7.37	6.39	7.01		16.28	18.45	12.74	16.87	15.40	14.83	
11. % S-Init. Directives	18.69	13.35	22.45	15.32	27.23	33.60		16.05	14.41	13.49	14.67	15.11	12.46	
12. % T-Init. Questions	33.21	20.53	33.95	17.71	36.26	21.53		34.70	21.56	40.62	24.07	34.27	19.89	
13. % T-Init. Comments	5.60	6.44	5.20	5.58	2.67	3.06	.09	6.21	9.74	5.39	5.70	3.66	4.23	
14. % T-Init. Academic Contacts	47.37	24.73	50.54	22.90	55.20	30.99		46.51	24.32	47.93	28.43	42.49	22.27	
15. % T-Init. Behav. Contacts	4.98	4.45	6.04	5.82	5.07	5.28		4.66	6.06	4.43	7.31	3.86	3.80	
16. % S-Init. Questions	5.16	5.34	4.11	5.63	3.33	3.53		9.38	12.76	6.51	8.15	8.99	11.15	
17. % S-Init. Comments	3.75	4.75	4.41	4.49	3.09	4.08		6.91	8.59	6.23	10.30	6.41	7.00	
18. % S-Init. Ac. Contacts	6.16	6.19	6.16	6.00	4.84	5.58		10.39	10.56	8.47	11.73	11.35	11.39	
19. % Student-Init. Proc. Cont.	2.62	3.69	2.51	2.59	1.46	2.09		5.89	9.93	4.27	8.02	3.66	4.73	
20. #/hr. Contacts Praise*	1.31	2.74	.86	1.69	.66	1.56		..	..	..	..	..	..	
21. #/hr. Contacts Criticism*	1.85	2.52	2.66	3.23	2.51	3.22		..	..	..	..	..	..	
22. % Contacts T-Init.	88.20	9.46	89.37	7.93	88.43	18.39		78.82	18.38	85.21	14.74	77.89	18.93	

VARIABLE (READING)	GRADE 2 SMALL (n=36)		REGULAR (n=24)		AIDE (n=26)		SIG.	GRADE 3 SMALL (n=37)		REGULAR (n=21)		AIDE (n=26)		SIG.
	M	SD	M	SD	M	SD	SD	M	SD	M	SD	M	SD	SD
23. % Contacts Academic	80.96	13.73	82.60	11.68	81.79	15.21		80.48	16.82	80.80	18.87	77.20	19.20	
24. % Contacts Behavioral	7.47	6.07	7.82	6.93	5.30	5.87		6.06	7.78	6.58	10.09	5.67	5.31	
25. % Contacts Directives	29.71	18.17	32.91	13.12	33.66	21.83		22.69	18.41	18.17	16.29	22.37	17.07	
26. % Contacts Questions	56.18	19.51	52.99	16.71	57.73	23.92		61.23	19.65	68.08	20.36	63.04	18.03	
27. % Contacts Praise *	2.13	4.10	1.23	2.52	.92	2.33		..	..	..	..	..	..	
28. % Contacts Criticism *	3.34	4.56	3.74	4.39	3.33	4.36		..	..	..	..	..	..	
29. % Time in Acad. Activities	97.27	4.24	98.13	3.16	97.46	2.84		95.63	5.29	89.64	11.56	93.81	7.75	.03
30. % Students Def. On-Task	88.78	7.87	83.50	10.62	90.29	8.67	.02	88.36	11.92	88.21	9.37	84.82	15.45	
31. % Students Prob. On-Task	3.92	3.10	5.25	3.47	2.62	2.05	.01	5.46	7.67	3.32	3.49	5.42	7.00	
32. % Students Off-Task	5.33	5.52	8.40	6.69	5.99	7.35		4.97	6.80	6.00	6.66	6.90	6.88	
33. % Students Waiting	1.58	2.16	2.80	3.86	1.02	2.31	.07	1.84	4.66	1.70	2.28	2.86	5.03	

* Data for this variable not collected for Grade 3.

TABLE VI-7. Means and Standard Deviations for Grade 2 and Grade 3 Class Types Aggregated Across Training Conditions. Math Variables

VARIABLE (MATH)	GRADE 2 SMALL M	SMALL SD	REGULAR M	REGULAR SD	AIDE M	AIDE SD	SIG.	GRADE 3 SMALL M	SMALL SD	REGULAR M	REGULAR SD	AIDE M	AIDE SD	SIG.
1. Avg. Time in Subject	38.97	13.84	38.15	11.85	32.72	9.23		40.65	15.09	36.32	17.59	42.92	9.22	
2. No. of Small Groups	.22	.96	.08	.41	.31	1.57		.24	.49	.11	.32	.31	.55	
3. Content Development	33.38	17.63	33.51	18.02	35.80	20.84		30.94	13.97	34.76	17.26	30.77	17.50	
4. Assignment Directions	3.57	6.62	3.47	5.61	4.34	9.23		3.52	5.31	4.50	5.49	3.62	6.19	
5. Independent Seatwork	15.72	15.66	12.32	11.83	13.96	15.85		16.57	15.03	13.77	15.79	16.15	13.64	
6. Small Group Instruction	1.00	3.73	.00	.00	.00	.00		.44	2.66	1.13	4.92	.00	.00	
7. Testing	2.71	10.87	4.81	12.13	3.17	12.40		1.07	3.55	.00	.00	.66	2.35	
8. % Individual Contacts	69.12	42.92	66.48	36.30	74.22	47.80		68.85	44.40	70.21	37.08	60.14	34.61	
9. % T-Init. Contacts	59.30	39.95	56.57	30.60	65.32	44.09		52.44	30.63	63.36	34.18	50.46	28.47	
10. % S-Init. Contacts	9.76	10.50	8.21	9.54	8.66	10.89		16.40	27.71	6.85	6.71	9.68	9.97	
11. % S-Init. Directives	15.85	14.49	18.35	14.08	23.47	28.05		15.13	13.84	12.73	16.64	20.37	25.50	
12. % T-Init. Questions	30.88	33.61	28.94	23.12	36.47	30.13		27.31	18.61	45.48	19.09	24.10	16.37	.0004
13. % T-Init. Comments	9.51	9.80	9.31	10.81	5.39	5.11		10.00	15.65	5.15	6.45	5.98	6.48	
14. % T-Init. Academic Contacts	44.33	33.12	43.04	26.49	51.67	37.65		41.41	23.81	53.50	27.95	34.72	20.19	.04
15. % T-Init. Behav. Contacts	5.82	8.16	7.86	7.85	5.58	6.09		4.91	8.07	3.52	5.95	4.34	6.44	
16. % S-Init. Questions	6.46	5.83	5.19	8.94	6.46	8.84		12.36	18.32	5.07	5.10	6.98	7.80	
17. % S-Init. Comments	3.16	5.82	3.04	3.01	2.30	3.74		4.04	6.37	1.79	2.53	2.70	4.31	
18. % S-Init. Ac. Contacts	7.02	7.86	4.78	5.34	6.53	9.17		10.54	14.62	5.44	5.71	5.39	6.11	
19. % Student-Init. Proc. Cont.	2.73	4.25	3.43	4.84	2.32	3.66		5.86	9.51	1.42	2.38	4.29	5.56	.10
20. #/hr. Contacts Praise *	1.43	3.10	3.93	8.84	1.13	2.30		..	..	..	..	..	..	
21. #/hr. Contacts Criticism *	2.06	2.82	1.73	2.28	1.89	3.17		..	..	..	..	..	..	
22. % Contacts T-Init.	77.70	25.89	87.94	13.24	80.69	26.01		80.00	21.72	90.99	9.06	86.97	11.43	.05

VARIABLE (MATH)	GRADE 2 SMALL M	SMALL SD	REGULAR M	REGULAR SD	AIDE M	AIDE SD	SIG.	GRADE 3 SMALL M	SMALL SD	REGULAR M	REGULAR SD	AIDE M	AIDE SD	SIG.
23. % Contacts Academic	71.48	25.02	74.05	18.38	77.19	20.70		78.95	21.57	86.61	11.90	69.92	19.98	.02
24. % Contacts Behavioral	7.77	9.28	11.38	8.30	7.17	7.60		6.13	9.14	3.73	5.00	8.23	9.92	
25. % Contacts Directives	27.28	22.73	30.01	17.27	28.85	23.49		25.29	20.09	16.95	18.21	30.57	26.56	
26. % Contacts Questions	55.18	24.89	50.64	19.81	59.82	25.03		57.09	24.00	75.38	20.32	55.17	27.08	.01
27. % Contacts Praise *	2.62	5.48	6.15	12.22	2.26	5.37		..	..	..	..	..	..	
28. % Contacts Criticism *	3.86	5.71	3.53	5.13	2.80	5.23		..	..	..	..	..	..	
29. % Time in Academ. Acts.	97.28	5.97	96.99	8.41	99.18	2.45		95.56	7.73	93.69	14.28	94.08	8.78	
30. % Students Def. On-Task	88.76	10.55	83.99	11.46	87.91	10.52		88.44	13.68	83.75	19.22	85.27	15.05	
31. % Students Prob. On-Task	3.62	4.80	3.80	2.50	3.33	2.78		4.07	7.22	2.74	3.05	2.94	3.60	
32. % Students Off-Task	4.29	7.13	4.93	5.77	5.08	8.54		5.95	9.38	7.39	12.06	5.76	6.89	
33. % Students Waiting	3.43	5.83	7.33	9.06	4.86	6.96		1.57	3.92	6.11	12.40	5.74	9.79	.08

* Data for this variable not collected for Grade 3.

TABLE VI-8. Means and Standard Deviations for Classroom Rating, Reading and Math Variables Disaggregated by Training Condition and Class Type Variables with Significant Interactions ($p < .10$) Only

VARIABLE	SMALL				REGULAR				AIDE			
	n=23		n=13		n=14		n=10		n=17		n=9	
	Training		Comparison		Training		Comparison		Training		Comparison	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Classroom Rating Grade 2												
None												
Classroom Rating Grade 3												
17. Students Successful in Lessons	4.00	.74	4.29	.83	4.15	.80	3.88	.99	3.53	.94	4.22	.67
19. Rewards Good Performance	3.87	.97	4.14	.95	3.62	1.12	4.13	.99	3.65	1.06	4.00	1.41
Reading Grade 2												
11. % T-Init. Directives	18.14	14.17	19.68	12.25	21.25	16.17	24.13	14.72	19.04	12.20	42.71	53.08
25. % Contacts Directives	26.69	17.56	35.04	18.69	31.10	14.97	35.43	10.21	25.68	15.42	48.73	24.98
31. % Students Prob. On-Task*	3.97	3.40	3.83	2.62	6.65	3.30	3.29	2.79	2.82	2.10	2.24	2.01
33. % S Waiting*	1.58	2.11	1.58	2.34	3.72	4.50	1.51	2.40	.47	.70	2.06	3.71
Reading Grade 3												
7. Testing	.40	1.85	.69	2.57	1.27	2.96	.00	.00	.00	.00	.93	2.80
12. % T-Init. Questions	28.47	16.92	44.49	24.89	43.14	18.28	37.16	31.45	28.92	17.96	44.37	20.38
17. % S-Init. Comments	6.22	8.41	7.98	9.07	7.65	12.61	4.28	5.99	4.52	5.12	9.98	8.88
25. % Contacts Directives *	26.43	19.10	16.82	16.18	12.46	14.58	26.02	16.03	25.75	19.22	16.00	10.06
26. % Contacts Questions	57.72	19.63	66.76	19.06	72.81	20.60	61.57	19.39	60.53	19.73	67.80	14.08
29. % Time in Ac. Activities	95.46	8.20	95.74	7.18	95.81	12.13	90.06	17.83	93.82	7.97	94.57	10.64
Math Grade 2												
16. % S-Init. Questions	6.08	5.84	7.13	5.98	3.49	5.46	7.56	12.27	7.93	9.18	3.67	7.90
19. % S-Init. Proc. Contacts*	1.75	2.02	4.47	6.33	2.64	3.76	4.54	6.09	3.56	4.03	.00	.00
23. % Contacts Academic	67.70	27.15	78.17	19.96	76.21	19.53	71.04	17.17	70.85	22.09	89.18	10.76
28. % Contacts Criticism	4.76	6.50	2.25	3.66	2.56	4.06	4.90	6.32	3.55	6.20	1.37	2.25
Math Grade 3												
6. Small Group Instruction	.00	.00	1.15	4.32	1.79	6.19	.00	.00	.00	.00	.00	.00

* $p < .05$

TABLE VI-9. Summary of Training Effects, Class Type Effects, and Interaction Effects For All Classroom Rating Variables

	Training Effect		Class Type Effect		Interaction Effect	
	G2	G3	G2	G3	G2	G3
1. Suitable Traffic Patterns	.05		Sml>Aide*			
2. Good Visibility	.05	.10	Sml>Reg* Sml>Aide*			
3. Describes Objectives Clearly	.01					
4. Materials are Ready	.02					
5. Clear Directions for Assignments						
6. Individualized Assignments						
7. Provides, Seeks Rationales						
8. Appropriate Pacing of Lesson		.05				
9. Clear Explanations, Presentations	.03					
10. Monitors Student Understanding						
11. Enforces Work Standards						
12. Efficient Administrative Routines	.05					
13. Appropriate General Procedures	.02					
14. Efficient Small Group Procedures						
15. Routines for Academic Work						
16. Considers Attention Spans						
17. Successful Students					.06	
18. Actions Related to Students' Interest	.09					
19. Rewards Good Performance						.06
20. Consistent						
21. Effective Monitoring						

22. Efficient Transitions	.04	.07
23. Disruptive Behavior		
24. Stopped Quickly		
25. Ignored		
26. Inappropriate Behavior		
27. Stopped Quickly		
28. Ignored	.07	
29. Task-Oriented Focus	.08	.01
30. Relaxed, Pleasant Atmosphere		.06
31. Avoidance Behavior During Seatwork	.05	

TABLE VI-10. Summary of Training Effects, Class Type Effects, and Interaction Effects For All Reading and Math Variables

*p < .05

	Training Effects				Class Type Effects				Interaction Effects			
	Reading		Math		Reading		Math		Reading		Math	
	G2	G3	G2	G3	G2	G3	G2	G3	G2	G3	G2	G3
1. Average Time in Subject												
2. N. of Small Groups												
3. Content Development	.006		.04									
4. Assignment Directions												
5. Independent Seatwork			.06									
6. Small Group Instruction	.008										.10	
7. Testing										.09		
8. % Individual Contacts												
9. % Teacher-Initiated Contacts												
10. % Student-Initiated Contacts												
11. % Teacher-Initiated Directives	.09									.10		
12. % Teacher-Initiated Questions		.03					Sml<Reg*** Aide<Reg***			.08		
13. % Teacher-Initiated Comments	.009				Aide<Reg Aide<Sml							
14. % Teacher-Initiated Academic Contacts							Sml<Reg* Aide<Reg*					
15. % Teacher-Initiated Behav. Contacts												
16. % Student-Initiated Questions											.07	
17. % Student-Initiated Comments										.10		
18. % Student-Initiated Academic Contacts												
19. % Student-Initiated Procedural Cont.							Sml>Reg				.03	
20. #/hr. Contacts Res. in Praise												

21. #/hr. Contacts Res. in Criticism								
22. % Contacts Teacher-Initiated					Sml<Reg°			
23. % Contacts Academic	.06	.10			Aide<Reg°			.07
24. % Contacts Behavioral								
25. % Contacts Directives	.004						.07	.02
26. % Contacts Questions	.05				Sml<Reg°° Aide<Reg°°			.08
27. % Contacts Res. in Praise								
28. % Contacts Res. in Criticism	.07							.10
29. % Time in Acad. Activities				Sml>Reg° Aide>Reg°				.10
30. % Students Definitely On-Task			Sml>Reg° Aide>Reg°					
31. % Students Probably On-Task		.10	.03	Aide<Reg°° Sml<Reg°°				.04
32. % Students Off-Task								
33. % Students Waiting				Aide<Reg Sml<Reg	Sml<Reg		.02	

°p<.05
°°p<.01
°°°p<.001

Table VI-11

Mean Stanford Achievement Test Scale Scores, for Trained and Untrained Teachers, by Class Type, Second and Third Grades.

Test and Training	Class Type			
Reading	**Small**	**Regular**	**Regular/Aide**	**Total**
Second Grade				
Trained	595.6	585.2	591.4	591.2
Not Trained	593.9	583.9	585.3	587.5
Math				
Trained	593.5	586.5	582.3	587.1
Not Trained	590.2	580.5	580.8	583.5
Third Grade				
Reading	**Small**	**Regular**	**Regular/Aide**	**Total**
Trained				
Not Trained	621.3	611.4	612.7	614.9
Math				
Trained	622.2	617.1	617.7	619.0
Not Trained	622.8	615.1	615.6	617.6

Teachers with out-of-range classes are excluded. The total included is 303 in third grade, and 308 in second grade.

Table VI-12

Mean Stanford Achievement Test Gain Scores, Reading and Math for Trained and Untrained Teachers, by Class Type, Second and Third Grades.

Test and Training	Class Type			
Second Grade				
Reading	**Small**	**Regular**	**Regular/Aide**	**Total**
Trained	53.1	59.1	63.6	58.6
Untrained	57.9	58.0	58.6	58.2
Math				
Trained	47.3	47.2	45.2	46.5
Untrained	44.6	43.5	47.3	45.3
Third Grade				
Reading	**Small**	**Regular**	**Regular/Aide**	**Total**
Trained	23.8	30.2	24.3	25.7
Untrained	27.4	27.5	27.4	27.4
Math				
Trained	28.7	30.6	35.4	31.9
Untrained	32.6	34.6	34.8	34.1

Teachers with out-of-range classes are excluded from the analysis.