1999-2000 EVALUATION RESULTS OF THE STUDENT ACHIEVEMENT GUARANTEE IN EDUCATION (SAGE) PROGRAM

Submitted by the SAGE Evaluation Team
School of Education
University of Wisconsin–Milwaukee

Alex Molnar
Philip Smith
John Zahorik

Research Assistants

Karen Ehrle
Anke Halbach
Barbara Kuehl

DECEMBER 2000

CERAI-00-34

For further information, contact Alex Molnar, School of Education
University of Wisconsin–Milwaukee, P.O. Box 413, Milwaukee, WI 53201, (414) 229-2220.

This document is available on the SAGE Website:
http://www.uwm.edu/Dept/CERAI/sage.html

TABLE OF CONTENTS

INTRODUCTION
SAGE Program……3

Goals of SAGE Evaluation……..4

Class Size Research Background…….5

Overview of Findings from Previous Years ……...9

1999-2000 SAGE EVALUATION
Descriptions and Definitions…….14

Data Collection Instruments……..18

ANALYSES OF STUDENT ACHIEVEMENT OUTCOMES 1999-2000
Methods Introduction…….……22

SAGE School/Classroom vs. Comparison School/Classroom Analyses ……..26

Additional Analyses…….……..52

ANALYSES OF SAGE CLASSROOMS AND SCHOOLS 1999-2000 …….57

Qualitative Analysis Results……...……...63

Teacher Behavior and Student Achievement Analysis ……..…83

Discussion and Implications ……..……85

Teacher and Principal Questionnaires ……..86

MAJOR FINDINGS AND DISCUSSION 1999-2000…….99

REFERENCES ……...104

APPENDICES…...……..…….……...106

INTRODUCTION

SAGE Program

The Student Achievement Guarantee in Education (SAGE) evaluation is being conducted under contract with the Department of Public Instruction by the School of Education at the University of Wisconsin–Milwaukee. The purpose of the SAGE evaluation is to determine the effectiveness of the Student Achievement Guarantee in Education (SAGE) program in promoting academic achievement of students in kindergarten through third grade classrooms in schools serving low-income children.

The 1995 SAGE statute [s. 118.43] required participating schools to (1) reduce class size to 15 in kindergarten and grade one in 1996–97, grades kindergarten through two in 1997–98, and grades kindergarten through three in 1998–99 to 2000–2001; (2) stay open from early in the morning to late in the day and collaborate with community organizations to provide educational, recreational, community, and social services (i.e., the "lighted schoolhouse"); (3) provide a rigorous academic curriculum to improve academic achievement; and (4) establish staff development and accountability mechanisms.

The SAGE evaluation involves the 30 schools in 21 school districts that launched the SAGE program in 1996-97 in kindergarten and first grade. Second grade was added in 1997-98, and third grade in 1998-99. The SAGE evaluation compares SAGE schools to a group of 14-17 non-SAGE Comparison schools located in SAGE districts. The results of the 1996-97, 1997-98, 1998-99, and 1999-00 evaluations are generally consistent with Tennessee's Student Teacher Achievement Ratio (STAR) Project (1985-1989), the largest and best-controlled study on class size reduction to this point. It is worth noting two significant design differences in the Tennessee and SAGE studies. First, the Tennessee STAR Project used a true experimental design. The

SAGE project uses a quasi-experimental design. The SAGE project evaluation uses naturally occurring classrooms while STAR employed random assignment of students to classroom types which were held constant for the duration of the study. Second, the SAGE evaluation uses a control or comparison group for the purpose of assessing the impact of SAGE interventions.

Goals of SAGE Evaluation

The SAGE evaluation is intended to determine the impact that the four interventions of the SAGE program have on student achievement. To ascertain and to explain this impact, the evaluation addresses the following questions:

SAGE vs. Comparison School – Achievement Outcome Questions

1. What are the achievement levels of SAGE classrooms compared to achievement levels of classrooms in selected Comparison schools?
2. Does participation in a SAGE classroom have a differential impact on the achievement of minority students and white students?
3. Does the impact on achievement of participation in a SAGE classroom change from year to year as students progress from first through third grade?
4. Is the socioeconomic status (as measured by participation in the school lunch program) of SAGE participants related to individual achievement gains in first through third grade?
5. Do different types of SAGE classrooms (e.g. one teacher with 15 students vs. two teachers with 30 students) have different impacts on student achievement?

SAGE Schools – Classroom and School Questions

1. What are the instructional characteristics of SAGE classrooms?
2. How are SAGE classrooms organized?
3. Does the degree of congruence between SAGE classroom curricula and national professional curriculum standards in reading/language arts and mathematics correlate with the achievement levels in SAGE classrooms?
4. Does participation in the SAGE program result in an increase in the number or change in the type of school and/or community activities housed in the school before and after school hours?
5. What is the structure and focus of professional development activities in SAGE schools?
6. Does the number of years of teaching experience of SAGE teachers correlate with student achievement?

Class Size Research Background

The principal SAGE intervention is a reduction in class size to 15:1 in kindergarten through third grade classrooms. Class size reduction in the early elementary grades has become an increasingly popular issue for policymakers and researchers in recent years (Grissmer, 1999; U. S. Department of Education and the Laboratory for Student Success, 1999). Class size reduction has been credited with more learning opportunities for students, increased opportunities for teachers to meet children's individual needs, and less time spent on discipline problems. Parents and teachers like the idea and policymakers are embracing it. Several states, among them California, Florida, Indiana, Nevada, Tennessee, and Wisconsin have launched class size reduction efforts (Pardini, 1998; Viadero, 1998).

Decades of research on class size reduction have shown small achievement gains for students when, for example, class size was reduced from 25 to 20 students. In general, though, reductions of just a few students per class do not seem to significantly raise academic achievement. However, in the late seventies, an analysis by Glass suggested that larger reductions produced greater achievement gains (Glass, 1978; Pate-Bain, Achilles, Boyd-Zaharias, & McKenna, 1992).

A statewide experiment in Tennessee, the largest and best-controlled study on class size reduction to this point, assigned kindergarten students on a random basis to classes of 15, 25 with a teacher and an aide, or 25 with a teacher. The same configurations were maintained through third grade. Tennessee's Project STAR (Student/Teacher Achievement Ratio) analyzed student achievement in relation to class size over a four-year period (1985-1989). The project included 17 inner-city, 16 suburban, 8 urban, and 39 rural schools. Findings showed that students in the smaller classes scored higher on the Stanford Achievement Test and on the Basic Skills First

(BSF) Test in all four years (K-3) and in all locations. The greatest gains on the Stanford Test were made by inner-city small classes. While all students in small classes benefited, disadvantaged minority students seemed to benefit more from smaller class sizes than their peers (Pate-Bain, Achilles, Boyd-Zaharias, & McKenna, 1992). A further analysis of the use of teacher aides to reduce class size indicates that none of the differences in aide/regular classes achievement data was statistically significant, nor did teachers report advantages with regard to student behavior or reduction of teaching burdens (Finn, Gerber, & Farber 2000).

Studies such as STAR and SAGE can provide crucial information for policymakers. For example, a review of the research literature conducted by Bingham (1993) on white-minority achievement gap reduction and small class size, which included the STAR Project, suggests that small class sizes in the early grades represent an effective strategy to reduce the achievement gap. Bingham proposes that class size reduction may offer an early intervention strategy that serves to prevent rather than to reduce the achievement gap between white and minority students.

Further support for small classes in lower elementary grades is produced by the Lasting Benefits Study (LBS). The LBS tracked students who participated in Project STAR in order to determine whether achievement advantages of students from small classes were maintained after students returned to regular-sized classes in fourth grade. Data from 1990-1994 indicate that students who were originally in smaller classes continued to perform better than their peers from classes of 25 with or without a teacher's aide (Mosteller, 1995). Moreover, results from follow-up studies that have indicated that middle school students who attended STAR small classes were from 4 to 8 months ahead when compared to peers without the small class benefit in early elementary grades. Attendance in STAR small classes also appears to have beneficial effects on

minority students' college entrance exams and college test-taking results (Boyd-Zaharias & Pate-Bain, 2000).

Critics of class size reduction (CSR) say that it is one of the most expensive reforms in education and question whether the benefits are worth the cost. Overall achievement gains are minimal, they argue, and the investment required for producing the desired results of class-size reductions could be better spent elsewhere in education (Hanushek, 1999). Hruz (1998, 1999) argues that CSR policies compete with other educational reform measures, require a considerable commitment of funds, and their implementation can have a significant impact on the availability of qualified teachers.

Disagreements about the extent of benefits derived from CSR efforts are often grounded in mistaking CSR with ratios that compare number of students to number of adults working with students in a school. Hanushek's calculations, as other researchers have noted, were based, in many instances, on the ratio of students to staff and often included librarians and special education teachers who do not contribute to the actual reduction of number of students in a classroom (Viadero, 1998). Moreover, Krueger (2000), in a re-analysis of small class size findings, points out that Hanushek used estimates and a disproportionate weighting scheme which helped lead to findings of minimal positive effects of class size interventions. As Krueger indicates, Hanushek's approach to selection of estimates resulted in "the oversampling from studies with lower performance" (p. 19). Further, many of Hanushek's estimates were taken from studies not initially designed to study the effects of class size *per se* but some other component of education. Moreover, Hanushek's analyses of the relationship between the amount of money spent and student achievement outcome did not include the STAR Study.

Indeed, questions of class size and student performance involve the study of how resources can be most effectively allocated to produce desired outcomes. In education, as in other areas of society, the relationship between schooling inputs and schooling outputs is of interest. Wenglinsky (1997) studied the relationship between spending and student achievement by analyzing data from three separate sources: The National Assessment of Educational Progress, the Common Core of Data, and the Teacher's Cost Index of the National Center for Education Statistics. Wenglinsky's research suggests that increased spending to reduce class size has a direct positive effect on mathematics achievement for fourth grade students.

Grissmer (1999) notes that ultimately our confidence in measurements of small class size effects will arise from knowledge of the specific processes inside classrooms. Class size effects are better understood if we can discern the processes used in these classrooms that affect student development and achievement. Grissmer advises researchers to address fundamental questions about the nature of evidence in small class size studies. For policy decisions, he notes that "the seeming transparency of experimental data to policymakers should not be the deciding factor in their policy judgments" (p. 93). Grissmer further observes that having both experimental and non-experimental evidence is rare in the educational research community, something the SAGE evaluation project has been able to achieve. The design of the SAGE evaluation project utilized data from student tests as well as processes in SAGE classrooms. Analyzing processes inside small classrooms with regard to teacher and student behavior has been a crucial part of the SAGE evaluation project and that analysis is reported in the evaluation results along with student achievement data.

Grissmer (1999), from an analysis of small class size research, concludes that a key contribution of experimental data can be their guiding role in non-experimental studies to

develop a theory of classroom teacher and student behavior that explains higher student achievement. In Wisconsin, the SAGE evaluation addresses need.

Overview of Findings from Previous Years (1996-97, 1997-98, and 1998-99)

Achievement Outcome Findings 1996-99

To measure academic achievement, first grade students in SAGE schools and in a group of Comparison schools were tested in October 1996, and again in May 1997, using the Comprehensive Test of Basic Skills (CTBS) Complete Battery, Terra Nova edition, Level 10.

After one year, students in SAGE first grade classrooms scored higher on CTBS tests than first grade students in Comparison schools. As a group, SAGE students scored significantly higher on the post-test in reading, language arts, and mathematics sub-tests of the CTBS. The total score of SAGE students was also significantly higher than the total score of comparison group students. The achievement advantage associated with participation in the SAGE program was revealed both in the analysis of individual student scores and in the analysis of averaged classroom scores.

At the individual level of analysis, after controlling for pre-test scores, socioeconomic status (SES) as defined by eligibility for subsidized lunch, absenteeism, and race and ethnicity, SAGE first grade students scored higher than Comparison school first grade students on the CTBS post-test in reading, language arts, mathematics and total score. The results were statistically significant for all but the reading scores. At the classroom level, the post-test performance of SAGE first grade students was 4 scale score points higher in language arts, 4.3 scale score points higher in reading, 4.6 scale score points higher in mathematics, and 4.6 scale

score points higher in the total test score than Comparison school students. Each of these findings was statistically significant.

After adjusting for individual pre-test results, socioeconomic status (SES) as defined by eligibility for subsidized lunch, and student attendance, participation in SAGE shows a statistically significant advantage of 6.4 scale score points in the total score and 8.1 scale score points on the mathematics sub-test.

The classroom level data on the averaged performance of students in 1996-97 and 1997-98 SAGE classrooms suggested that the lower student-teacher ratio in SAGE classrooms mitigated the negative achievement consequences of poverty. SAGE classrooms achieved at a higher level than Comparison school classrooms despite the fact that, as a group, SAGE classrooms enrolled more students who were eligible for subsidized lunch. Furthermore, after adjusting for individual pre-test results and SES as defined by lunch status and student attendance, the post-test scale score advantage increased to 9.8 for SAGE first grade classrooms. The advantage was 7.1 on the reading sub-test, 9.0 on the language arts sub-test, and 12.3 on the mathematics sub-test. These results were all statistically significant.

Second grade classrooms were included in 1997-98, and test results were similar to those found for 1996-97 first graders. The 1997-98 results suggest that the positive effects of the SAGE program are maintained, but not increased in second grade.

Third grade classrooms were included in 1998-99, and test results suggest that statistically significant positive effects of SAGE, which occurred in first grade, were maintained in second and third grade. In 1998-99, African American SAGE students performed significantly higher on every subtest and total score over African American Comparison students on the third grade test.

Analyses of test results at the class level suggest that students in smaller classrooms tend to score significantly higher in language arts, mathematics, and reading as well as total score after adjusting for individual pre-test results, socioeconomic status, and attendance. In other words, classrooms with fewer students are more likely to have higher class average achievement scores and are more likely to contribute to closing the achievement gap between African American and white students than classrooms with a higher number of students.

School and Classroom Findings 1996-99

To more fully understand the impact of the SAGE program, it is important to understand how SAGE schools structure classrooms and implement 1) reduced student-teacher ratio, 2) rigorous curriculum, 3) staff development, and 4) lighted schoolhouse. Together, the information from all facets of the SAGE program provides a description of life in SAGE classrooms and schools and a more complete picture of the impact of the SAGE program on student performance.

School Level Findings

The Teacher Questionnaire and Principal Interviews, both completed in May 1997 and May 1998, were the sources of data regarding *rigorous curriculum* implementation. The Teacher Questionnaire contained a section on classroom curriculum designed to determine the congruence of SAGE classroom curricula with professional curriculum standards. First grade and second grade teacher responses indicated that their reading/language arts curriculum and mathematics curriculum were quite congruent with professional standards. Principal responses to curriculum-related questions suggested that a rigorous curriculum included basic skills, problem solving, and higher-level thinking. Only a few principals seemed to believe that the curriculum of their school was rigorous in every aspect. However, most SAGE principals regarded substantial parts of their curriculum as strong. A section of the Teacher Questionnaire contained *staff development* questions. Teachers were asked about their individual level of

professional development as well as the extent to which their school district provided staff development programs. About 60 percent indicated that they had a written development plan and it was determined by the teachers themselves. Data regarding implementation of *lighted schoolhouse* activities for 1996-97 and 1997-98 were obtained from the Principal Interviews and year-end reports required by DPI. Principal Interview data suggested that SAGE schools took responsibility for the conception and operation of the lighted schoolhouse activities and that the number of lighthouse activities and number of participants in the activities had progressively increased.

In 1998-99, a similar teacher questionnaire was administered and also a principal questionnaire was used to gather data. Principals indicated increasing implementation in curriculum, staff development, and lighted schoolhouse activities. Teachers described individualized instruction, teacher enthusiasm, and student engagement as prominent features of their teaching and generally perceived curricula to be congruent with national standards. Slightly less than half of the teachers reported that they have a personal, written professional development plan. Regarding parent contact, teachers indicated that most of the communication occurs in conversations and phone calls, notes sent home, and parental visits to the school.

Classroom Level Findings

Data from 1996-97 and 1997-98 suggested that the main change that results from having a reduced size class is individualization. Teachers focus on individual learning needs through one-to-one, small groups, and total class teaching. This focus on individuals came about because teachers knew students better, had more time for teaching because of reduced need for discipline, and were more enthusiastic about teaching, all which resulted from having fewer students.

The type of instruction that students encountered in SAGE classrooms was predominantly teacher centered. Listening, practicing, receiving help, and answering accounted for the main portion of the learning that occurred. Although teachers indicated that student-centered activities such as creating, manipulating, and problem solving increased somewhat

because of reduced class size, student-centered teaching played only a supplemental role in most SAGE classrooms.

In regard to teaching and learning among the different types of SAGE classes and between grades one and two, few differences were observed. Three case studies of different types of SAGE schools were conducted in 1998-1999. The case studies provided portraits of the functioning of SAGE schools and classrooms. Findings at the classroom level illustrated the various ways teachers individualize instruction and engage in interactive exchanges with students.

1999-2000 SAGE EVALUATION
Descriptions and Definitions

Schools

During 1999-00, the SAGE program evaluation was continued in 30 schools located in 21 school districts throughout the state, as shown in Table 1. In addition, the SAGE evaluation included data from 16 Comparison schools located in 11 SAGE school districts. The number of Comparison schools in 1999-00 reflects a reduction by one school from the previous year.

Table 1. SAGE Schools 1999-2000

SAGE DISTRICTS AND SCHOOLS	
DISTRICT School	**DISTRICT** School
ADAMS-FRIENDSHIP AREA Adams Elementary	**MENOMONIE INDIAN** Keshena Primary
BELOIT Robinson Elementary	**MENOMONEE AREA** River Heights Elementary
CUDAHY Parkview Elementary	**MILWAUKEE PUBLIC SCHOOLS** Carleton Elementary Fairview Elementary Longfellow Elementary Maple Tree Elementary Maryland Avenue Elementary Sherman Elementary Wisconsin Conservatory
GILMAN Gilman Elementary	
GLIDDEN Glidden Elementary	
GREEN BAY AREA Jefferson Elementary	
JANESVILLE Wilson Elementary	**PRENTICE** Ogema Elementary Tripoli Elementary
KENOSHA Durkee Elementary	**SIREN** Siren Elementary
LAC DU FLAMBEAU #1 Lac Du Flambeau Elementary	**PORT WING** South Shore Elementary
LACROSSE Franklin Elementary Hamilton Elementary	**SUPERIOR** Blaine Elementary Cooper Elementary
LAONA Robinson Elementary	**SURING** Mountain Elementary
MADISON METROPOLITAN Glendale Elementary	**WEBSTER** Webster Elementary

Students

In 1999-00, the SAGE evaluation involved a total of 5,283 students, 2627 in second grade and 2,656 in third grade. The characteristics of students in SAGE and Comparison schools are displayed in Table 2.

Table 2. Characteristics of SAGE and Comparison Students 1996-97, 1997-98, 1998-99, and 1999-00

Characteristic	**Percent of Students* SAGE**				**Percent of Students* Comparison**			
	96-97	97-98	98-99	**99-00**	96-97	97-98	98-99	**99-00**
Gender								
Female	48.6	48.4	48.6	48.6	49.4	48.5	48.7	48.2
Male	51.4	51.6	51.4	51.3	50.6	51.5	51.3	51.8
Race/Ethnicity								
African American	24.8	26.3	22.4	25.3	32.9	24.7	19.7	27.4
Asian	5.7	5.2	4.8	5.2	5.5	5.6	5.9	6.5
Hispanic	6.6	6.5	6.4	7.8	8.0	10.0	9.5	12.5
Native American	11.7	10.3	10.9	10.4	1.4	1.5	1.5	1.3
White	48.8	43.8	44.2	46.9	49.0	52.2	53.4	48.5
Other	1.6	2.0	1.8	1.4	2.7	2.3	2.3	2.4
Subsidized Lunch Eligibility								
Free	57.7	54.0	52.7	53.1	49.4	43.4	40.7	48.4
Reduced	10.9	10.6	11.5	12.3	9.9	8.9	10.4	11.2
Not Eligible	31.4	35.4	35.8	31.6	40.7	47.7	48.8	38.6
Repeating Grade	3.2	2.7	2.0	1.6	2.6	2.0	1.5	1.0
English as Second Language	8.2	7.9	7.5	7.0	4.9	6.4	6.7	9.2
Referred to M-Team	13.6	9.6	12.7	13.2	9.2	6.8	9.1	11.3
Exceptional Education Need	13.1	10.0	12.7	13.7	9.7	7.1	1.3	11.1

(*Percentages may not always total to 100% due to incomplete reports submitted by some schools.)

During the course of the 1999-00 school year, records were compiled on 5,283 students. Many students withdrew from SAGE and Comparison schools during the year, while others enrolled as new students. The number of students in SAGE and Comparison schools by grade and school year can be seen in Table 3.

Table 3. Number of Students in SAGE and Comparison Schools by Grade and School Year

	SAGE				COMPARISON*			
	1996-97	**1997-98**	**1998-99**	**1999-00**	**1996-97**	**1997-98**	**1998-99**	**1999-00**
Kindergarten	1494	1524	1416	NA	820	676	887	NA
First Grade	1723	1567	1525	NA	1001	985	983	NA
Second Grade	NA	1541	1446	1636	NA	868	1047	991
Third grade	NA	NA	1531	1611	NA	NA	1041	1045
Totals	3217	4632	5918	3247	1821	2529	3958	2036

*The number of Comparison schools participating in the study since 1996 has fluctuated from 14 to 17. Student numbers for Comparison schools reflect this fluctuation.

Table 4 illustrates the stability of student enrollment for SAGE and Comparison schools. Those students who remained in their schools for the entire year are labeled "ongoing." These data were obtained from student profiles completed by the schools. The percentage of ongoing SAGE students was higher than the percentage of ongoing Comparison students, and the percentage of students withdrawing was lower in SAGE Schools than in Comparison Schools. The percentages for new student enrollment were similar between SAGE and Comparison Schools.

Table 4. Enrollment Changes in SAGE and Comparison Schools by School Year (Percentage of Students)

	SAGE				COMPARISON			
	1996-97	**1997-98**	**1998-99**	**1999-00**	**1996-97**	**1997-98**	**1998-99**	**1999-00**
Ongoing	81.4	42.3	44.8	71.3	85.3	44.3	39.1	65.8
Withdrew	11.0	18.8	27.5	12.6*	8.9	18.6	20.2	17.2*
New	7.6	39.0	27.6	16.0	5.8	37.0	40.6	16.9

*Note: Thirty-five (35) students switched from a SAGE school to a comparison school (or vice versa). Only these students were counted as withdrawals.

Classrooms

SAGE schools reduced class size in several ways in order to meet statutory requirements. The SAGE legislation defines class size as "the number of pupils assigned to a regular classroom teacher." In practice, reduced class size has been interpreted as a 15:1 student-teacher ratio

(number of students per teacher in one classroom). Implementation occurs in the following ways:

- A *Regular* classroom refers to a classroom with one teacher. Most *regular* classrooms have 15 or fewer students, but a few exceed 15.
- A *2-Teacher Team* classroom is a class where two teachers work collaboratively to teach as many as 30 students.
- A *Shared-Space* classroom is a classroom that has been fitted with a temporary wall that creates two teaching spaces, each with one teacher and about 15 students.
- A *Floating Teacher* classroom is a room consisting of one teacher and about 30 students, except during reading, language arts, and mathematics instruction when another teacher joins the class to reduce the ratio to 15:1.

Three other types of classroom organization have also been utilized in the SAGE program, but to a limited extent. One type is the *Split Day* classroom consisting of 15 students and two teachers, one who teaches in the morning and one who teaches in the afternoon. Another type is the *3-Teacher Team* classroom where 45 students are taught collaboratively by three teachers. These two types of classroom organization were not used in 1999-00. A third type consists of a *full time and part time teacher* combination to reduce class size for part of the day. This type of classroom organization was used in 7 classrooms in 1999-00. The types of classroom organization utilized in 1999-00 are displayed in Table 5.

Table 5. Number of SAGE Classrooms by Type and Grade 1999-00

	Regular 15:1	**Team Taught 30:2**	**Floating Teacher**	**Shared Space**	**1 Full Time & 1 Part Time Teacher**
Grade 1	46	26	4	3	NA
Grade 2	61	15	NA	8	5
Grade 3	57	17	NA	7	2
Mixed	6	NA	NA	NA	NA

SAGE classes range in number of students from 3 to 41. A few SAGE classrooms exceed the 15:1 student-teacher ratio by a few students. The teacher to student ratio for SAGE and Comparison classrooms can be seen in Table 6.

Table 6. Student-Teacher Ratio for SAGE and Comparison Classrooms 1999-00 (number of classrooms)

	SAGE Classrooms		**Comparison Classrooms**	
Students Per Teacher	**Second Grade**	**Third Grade**	**Second Grade**	**Third Grade**
3-13 Students	35	22	2	0
14-16 Students	35	38	0	0
17+ Students	23	24	39	43
Average Class Size	15.98	17.94	22.05	22.91

Data Collection Instruments

To provide information about the processes and outcomes of the SAGE program for 1996-97, 1997-98, 1998-99, and 1999-00, a number of instruments were used as part of the evaluation.[1] A description of the test and non-test instruments used in 1996-97, 1997-98, 1998-99, and 1999-00 follows. The data collection instruments and the plan for their use throughout the evaluation are displayed in Table 7.

1. *Comprehensive Test of Basic Skills (CTBS)* (1996-97, 1997-98, 1998-99, 1999-00). The Comprehensive Test of Basic Skills (CTBS) complete Battery, Terra Nova edition, Level 10, was administered to first grade students in SAGE schools and Comparison schools in October 1996 and May 1997. In 1997-98, Level 10 was administered in October and Level 11 in May to first grade students and Level 12 to second grade students. In 1998-99, Level 10 was administered in October and Level

[1]See the *Evaluation Design Plan for the Student Achievement Guarantee in Education (SAGE) Program*, August 13, 1996, for complete details.

11 in May to first grade students, Level 12 to second grade students, and Level 13 to third grade students. The purpose of the first grade October administration of the CTBS was to obtain baseline measures of achievement for SAGE schools and Comparison schools. The complete battery includes sub-tests in reading, language arts, and mathematics. The CTBS was chosen as an achievement measure because it is derived from an Item Response Theory (IRT) model that allows comparison of performance across time. Moreover, it is one of a few instruments that attempts to minimize items biased against minorities and educationally disadvantaged students. Kindergarten students were not tested because of (1) concerns over the reliability and validity of standardized test results for kindergarten-aged children and (2) the view expressed by many kindergarten teachers that standardized tests would have a traumatizing effect on their students. The effects of SAGE on kindergarten students will be determined when they are tested as first-grade students the following year.

2. *Student Profiles* (1996-97, 1997-98, 1998-99, 1999-00). This instrument completed in October and May, provided demographic and other data on each SAGE school and Comparison school student.
3. *Classroom Organization Profile* (1996-97, 1997-98, 1998-99, 1999-00). Completed in October, this instrument was used to record how SAGE schools attained a 15:1 student-teacher ratio.
4. *Principal Interviews* (1996-97 and 1997-98). These end-of-year interviews elicited principals' descriptions and perceptions of effects of their schools' rigorous curriculum, lighted-schoolhouse activities, and staff development program, as well as an overall evaluation of the SAGE program. In 1998-99, principal interviews were conducted in the schools selected for case studies.
5. *Teacher Questionnaire* (1996-97, 1997-98, 1998-99, 1999-00). Administered in May, this instrument obtained teachers' descriptions and judgments of the effects of

SAGE on teaching, curriculum, family involvement, and professional development. It also was used to assess overall satisfaction with SAGE.

6. *Teacher Activity Log* (1996-97, 1997-98). This instrument required teachers to record classroom events concerning time use, grouping, content, and student learning activities for a typical day three times during the year.
7. *Student Participation Questionnaire* (1996-97, 1997-98). In both October and May, teachers used this instrument to assess each student's level of participation in classroom activities.
8. *Classroom Observations* (1996-97, 1997-98, 1998-99, 1999-00). A group of first grade, second and third grade classrooms representing the various types of 15:1 student-teacher ratios and a range of geographic areas was selected for qualitative observations to provide descriptions of classroom events.
9. *Teacher Interviews* (1996-97, 1997-98, 1998-99, 1999-00). Although in-depth teacher interviews were not part of the original SAGE evaluation design, they were added in 1997-98 because it became apparent that teachers had important stories to tell about their SAGE classroom experiences. The interviews deal with teachers' perceptions of the effects of SAGE on their teaching and on student learning.
10. *Principal Questionnaire* (1998-99, 1999-00). The Principal Questionnaire was administered to all SAGE principals in spring beginning in 1998-99. It asked them to rate and comment on teaching, rigorous curriculum, staff development, and lighted schoolhouse activities.
11. *Case Studies* (1998-99). Case studies of teaching in three schools, each representing a different type of SAGE class configuration, were conducted continuously throughout the school year in 1998-99. At grades one, two and three classrooms were observed in reading-language arts instruction and mathematics instruction and teachers were interviewed. Interviews with the principal and parents were also conducted.

12. *Classroom Studies* (1999-00). The teaching behaviors used by a group of highly effective, reduced class size first-grade SAGE teachers were compared to the teaching behaviors used by a group of less effective, reduced class size, first-grade SAGE teachers using qualitative research procedures.

Table 7. SAGE Data Collection by Grade Level, 1996–01

	1996–97	**1997–98**	**1998–99**	**1999-2000**	**2000-2001**
CTBS Fall, Spring Spring	1	1 2	1 2, 3	2, 3	3
Student Profiles Fall, Spring	1	1, 2	1, 2, 3	2, 3	3
Classroom Organization Profile Fall	1	1, 2	1, 2, 3	2, 3	3
Principal Interviews Spring	yes	yes	yes (selected)		
Teacher Questionnaire Spring	K, 1	K, 1, 2	K, 1, 2, 3	K, 1, 2, 3	K, 1, 2, 3
Teacher Activity Log Fall, Winter, Spring	K, 1	K, 1, 2	discontinued		
Student Participation Questionnaire Fall, Spring	K, 1	K, 1, 2	discontinued		
Classroom Observation Fall, Spring	1 (selected)	1, 2, (selected)	1, 2, 3 (selected)	1 (selected)	2, 3 (selected)
Teacher Interview Fall/Spring	1 (selected)	1, 2 (selected)	1, 2, 3 (selected)	1 (selected)	2,3 (selected)
Principal Questionnaire Spring	NA	NA	yes	yes	yes
School Case/Classroom Studies by grade level	NA	NA	1, 2, 3 (selected)	1 (selected)	2, 3 (selected)

ANALYSES OF STUDENT ACHIEVEMENT OUTCOMES 1999-00

Methods Introduction

Statistics Utilized

The 1999-00 SAGE evaluation design utilizes descriptive statistics and multivariate inferential statistics, including linear regression and hierarchical linear modeling. Descriptive statistics, including means and standard deviations, are incorporated into this report to provide a less complicated, more general analysis which the non-technical reader can use as a basis to interpret the findings. Regression analyses (at the individual level), specifically the use of ordinary least squares regression models, are employed frequently in the 1999-00 report. Regression models enable "control" variables to be entered in blocks with the variable of interest, i.e. the "SAGE/Comparison" variable entered last, thus isolating its effects from the other variables. Hierarchical linear modeling is pertinent to the SAGE evaluation because this technique permits a focus on the class effects of SAGE; that is, these analyses specifically assess classroom effects rather than those of individuals within the classroom. The classroom effects examined by this approach are of primary importance to the SAGE evaluation.

General Findings 1996-97

The quantitative findings from 1996-97show that first grade classrooms in SAGE schools scored higher on the CTBS Complete Battery, Terra Nova Level 10 than first grade students in Comparison schools. In general, when adjusted for pre-test scores, SAGE students scored significantly higher on the post-test in the areas of reading, language arts, and mathematics as well as total score. This means that not *all* SAGE students outperformed *all* Comparison students. At the individual level of analysis, after controlling for pre-test score, SES, attendance,

and race, SAGE first grade students as a group scored statistically significantly higher than Comparison school students on the CTBS post-test in the areas of language arts and mathematics as well as total score. At the class level of analysis, SAGE classrooms scored significantly higher in language arts, mathematics, and reading as well as total score after adjusting for individual pre-test results, SES, and attendance.

General Findings 1997-98

Analyses of the second year test data indicated that students in SAGE first grade classrooms achieved significantly higher scores than students in Comparison school classrooms in all tested areas: mathematics, reading, and language arts. The achievement advantage of students in SAGE first grade classrooms in 1996-97 was maintained in second grade in 1997-98. The advantage, however, did not appear to have increased significantly.

General Findings 1998-99

Analyses of the third year test data indicated that students in SAGE first grade classrooms again achieved significantly higher scores than students in Comparison school classrooms in all tested areas: mathematics, reading, language arts, and on the total scale score. The achievement advantage of students in SAGE first grade classrooms in 1997-98 was mirrored significantly in second grade in 1998-99, with the exception of reading. The achievement advantages of students in SAGE second grade classrooms in 1997-98 also continued to be statistically significant in 1998-99 third grade classrooms with the exception of language arts.

Score Metrics 1999-00

A brief discussion of the metrics reported in the 1999-00 SAGE evaluation is warranted. The SAGE report presents the findings using two metrics, scaled scores and normal curve equivalents. A scaled score provides a means for comparison across subjects or groups on a specific task or trait. A scaled score provides a common yardstick by which scores may be compared reasonably, subject to subject or group to group. The primary reason scaled scores are used in the SAGE quantitative analysis is to anchor the scores from test level to test level (level 10, 11, etc.) so that year-to-year results can be compared. Table 8 gives the nationwide Scale Score Descriptive Statistics for the test. Lowest and highest obtainable scale scores are provided for each level of each sub-test.

Table 8. Scale Score Descriptive Statistics

	Lowest Obtainable Scale Score	**Highest Obtainable Scale Score**
Reading		
Level 10 (first grade pre-test)	355	626
Level 11 (first grade post-test)	407	701
Level 12 (second grade)	423	722
Level 13 (third grade)	427	750
Mathematics		
Level 10 (first grade pre-test)	290	629
Level 11 (first grade post-test)	324	680
Level 12 (second grade)	347	720
Level 13 (third grade)	385	740
Language Arts		
Level 10 (first grade pre-test)	325	620
Level 11 (first grade post-test)	400	680
Level 12 (second grade)	424	706
Level 13 (third grade)	455	730

When comparing the scores to those of other individuals (or groups) to obtain meaning, we make a norm-referenced interpretation. Here the use of normal curve equivalents is useful.

A norm-referenced interpretation involves comparing a person's score with those of some relevant group of people. The normal curve equivalent scale ranges from 1 to 100 and thus provides a comparative index of the performance of an individual or group to the reference group. In this case, the reference group is the Terra Nova norm reference group (for norm referencing population data see CTB/McGraw-Hill, 1991). Normal curve equivalents are generally not good indicators of longitudinal progress, however. With these scores, the group average could remain at, for example 50, across pre-test and post-test with the reader erroneously concluding that no gain was made. Actually, if the focus group, in this example, did not "gain" more than the reference group, the score would remain constant although both groups scored higher.

Structure of 1999-00 Report

The descriptive analyses utilize both scale scores and normal curve equivalents. The inferential analyses (regressions and hierarchical linear models) utilize only scale scores. For the inferential tests, a significance level of .05 was used and significant results are denoted by an asterisk (*). SAGE versus Comparison analyses are divided into two major sections: Second Grade Results and Third Grade Results. The following are presented within each of these sections: (1) descriptive statistics (pre-test and post-test), (2) ordinary least squares regressions, (3) analyses of the scores of African American students, and (4) hierarchical linear models.

The quantitative section includes additional analyses looking at years of participation in SAGE, socio-economic status, and types of classroom. The quantitative section also includes "within SAGE" analyses for second and third grade students. SAGE student achievement is examined in relation to teacher experience, student participation, proximity to curriculum, and class organization.

SAGE School/Classroom vs. Comparison School/Classroom Analyses

Second Grade Results 1999-00

Descriptive Statistics

Valid Test Scores. Analyses were conducted to assess the impact of SAGE on the 1999-00 second grade CTBS Complete Battery, Terra Nova Level 12 post-test results. The number of second grade students with valid test scores for the Fall 1998 first grade pre-test, the Spring 1999 first grade post-test, or the Spring 2000 second grade test are presented in Table 9.

Table 9. Number of 1999-00 Second Grade Students with Valid First Grade Pre-Test or Post-Test Scores

	Fall 1998 First Grade Pre-test		**Spring 1999 First Grade Post-test**		**Second Grade Test 2000**	
	SAGE	**Comparison**	**SAGE**	**Comparison**	**SAGE**	**Comparison**
Reading	1004	581	1039	604	1338	804
Language Arts	1004	581	1039	604	1338	804
Mathematics	1005	580	1034	600	1340	797
Total	1000	579	1031	592	1318	786

In the analyses to follow, second grade post-test results are compared to the first grade pre-test, as well as the first grade post-test. Therefore, only those second grade students who took both the first grade pre-test and post-test, as well as the second grade test, were used in the 1999-00 second grade analysis. As would be expected, the number of second grade students having all three valid test scores was substantially less than the total number of students. Those students present in both the 1998-99 SAGE and Comparison first grade classrooms and the 1999-00 SAGE and Comparison second grade classrooms were termed “two-year persisters”. There

were 1520 two-year persisters, of which 969 were students in 30 SAGE schools and 551 were students in 16 Comparison schools. It should be noted that some persisters did not take all four CTBS tests. Table 10 shows the number of 1999-00 second grade persisters who took the test battery or one or more of the sub-tests.

Table 10. Number of 1999-00 Second Grade Two-Year Persisters with Valid First Grade Pre-Test and Post-Test Scores in Total Scale and/or one or more Subtests

	Fall 1998 First Grade Pre-test		**Spring 1999 First Grade Post-test**		**Second Grade Test 2000**	
	SAGE	**Comparison**	SAGE	**Comparison**	**SAGE**	**Comparison**
Reading	965	550	967	543	956	542
Language Arts	965	550	967	543	956	542
Mathematics	965	549	961	540	959	537
Total	961	548	959	532	946	528

Pre-Test (Baseline) Results. Both the first grade pre-test and the first grade post-test served as baseline measures. Tables 11, 12 and 13 provide descriptive statistics on the scale scores for SAGE and Comparison students from the first grade pre- and post-test as well as the second grade test. Given the longitudinal nature of the SAGE evaluation, scale scores serve as the primary measure of student achievement. To place the scale scores in context, national percentiles are also provided in the tables. For example, the mean or average reading scale score of 539.29 corresponds to a national percentile rank of 37.12. That is, the average first grade student in the SAGE evaluation pre-test scored as well on the CTBS Reading Scale as about 37 percent of students taking the test nationwide.

Table 11. Descriptive Statistics on CTBS First Grade Pre-Test and Post-Test (SAGE), 1999-00

	First Grade Pre-Test 1998	**First Grade Post-Test 1999**

	Scale Scores		Normal Curve Equivalent		Scale Scores		Normal Curve Equivalent	
	Mean	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**
Reading	539.29	37.12	47.56	20.64	583.33	39.72	53.95	20.52
Language Arts	541.71	39.97	49.50	21.33	589.27	39.41	57.14	19.22
Mathematics	501.40	37.51	47.27	18.44	543.41	38.20	53.93	18.65
Total	527.56	33.04	47.90	18.80	572.12	33.07	55.75	18.73

Table 12. Descriptive Statistics on CTBS First Grade Pre-Test and Post-Test (Comparison), 1999-00

	First Grade Pre-Test 1998				First Grade Post-Test 1999			
	Scale Scores		Normal Curve Equivalent		Scale Scores		Normal Curve Equivalent	
	Mean	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**
Reading	543.49	40.42	50.61	21.45	582.49	37.31	53.43	19.62
Language Arts	542.55	46.00	50.63	22.70	586.41	42.66	55.70	20.56
Mathematics	506.12	39.91	49.76	19.54	537.12	40.36	50.81	20.03
Total	530.83	35.97	50.15	19.75	568.86	35.06	53.92	19.72

Table 13. SAGE and Comparison Descriptive Statistics on CTBS Results for 1999-00 Second Grade Test

	Scale Scores				Normal Curve Equivalent			
	SAGE		Comparison		SAGE		Comparison	
	Mean	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D**	**Mean**	**S.D.**
Reading	612.35	33.68	610.22	33.91	52.18	17.63	51.94	17.89
Language Arts	616.00	40.87	609.15	38.51	52.71	20.50	49.21	19.18
Mathematics	579.17	41.05	566.62	40.25	53.37	20.36	47.10	20.82
Total	602.90	33.15	595.85	32.80	53.57	18.89	49.53	18.82

Difference of Means Test. The difference of means between SAGE and Comparison students on the 1999-00 second grade test can be seen in Table 14. SAGE students scored significantly higher in the total score and in every sub-test except for reading. The largest difference in means is found on the mathematics subtest, followed by the total score, and then language arts.

Table 14. Differences of Means Test 1999-00 Second Grade Test

	SAGE	Comparison	Difference
Reading	612.35	610.22	2.12
Language Arts	616.00	609.15	6.85*
Mathematics	579.17	566.62	12.55*
Total	602.90	595.85	7.05*
*significant at .05 level			

Statistically significant unadjusted gains are found in favor of SAGE students for each subtest and for total scale scores when the first grade pre-test was used as the baseline scores, as can be seen in Table 8. Unadjusted gains have not been corrected for pre-existing differences between groups on factors such as prior achievement, attendance, race and SES. The largest unadjusted gain for SAGE students from the first-grade pretest to the second-grade test was in mathematics. The smallest gain was in reading. When the first grade post-test was used as the baseline score, SAGE students continued to show significant gains, relative to Comparison students, in the math sub-scale and total score. This suggests that the statistically significant positive effects of SAGE occurred in the first grade. Furthermore, these positive effects were maintained and significantly increased in second grade in mathematics and the total score.

Table 15. SAGE and Comparison Unadjusted Gain for Second Graders, 1999-00

	From First Grade Pre-Test to Second Grade			From First Grade Post-Test to Second Grade		
SCALE SCORE	SAGE Gain	Comparison Gain	Gain Difference	SAGE Gain	Comparison Gain	Gain Difference
Reading	72.84	66.19	6.65*	28.75	27.55	1.20
Language Arts	73.88	65.78	8.10*	26.39	22.61	3.78
Mathematics	77.66	60.11	17.54*	35.48	29.36	6.12*
Total	74.83	64.24	10.59*	30.23	26.94	3.29*
*significant at .05 level						

Regression Analysis

As noted above, the unadjusted gains depicted in Table 15 above do not reflect group differences related to socioeconomic status, ethnicity, attendance and prior knowledge. In order to correct for group differences related to these factors, regression analysis was employed. Essentially, regression analysis allows for a statistical adjustment that "equalizes" the groups on factors where pre-existing differences exist.

Regression Models. The effect of the SAGE program on student achievement for second graders was tested through a series of ordinary least squares regression models for each sub-test and total scale score. Control variables were entered into the models in blocks, with the SAGE/Comparison student variable entered into the models last. Two different regressions were done for each sub-test and total scale score. The first regression used the first-grade pre-test as a predictor variable and the second regression used the first grade post-test as a predictor variable.

The first block of control variables included student scores on either the first grade pre-test or post-test and on attendance, measured as number of days absent, as reported by teachers in Spring 2000. Eligibility for subsidized lunch, as an indicator of family income, is also included in the first block of control variables. This variable is coded 1 if student is ineligible, 2 if student is eligible for reduced price lunch, and 3 if the student is eligible for free lunch (this variable is assumed to be interval level). The second block of control variables added dummy variables for race/ethnicity, coded 1 if a student was of a certain race/ethnicity, and 0 if not. Dummy variables were included for African Americans and whites. In the final block, a dummy variable for SAGE or Comparison school student was added. This variable is coded 0 if a student is from a Comparison school and 1 if a student is from a SAGE school.

Regression Results. Results of the regression analyses are presented in Tables 16 and 17. When the first grade pre-test was used as the predictor variable, membership in SAGE emerged as a significant predictor of student achievement on the total score and for all sub-tests. Similarly, the first grade post-test was a significant predictor of student achievement on total score and for all sub-tests. The magnitude of the effect of SAGE on student achievement, as denoted by the unstandardized regression coefficient, varies depending on the CTBS sub-test.

The largest effects of SAGE are found when the first grade mathematics pre-test is used to predict the second grade test (17.391). When all cases are analyzed, the goodness-of-fit of the models (as denoted by the adjusted R square statistic), ranged from .23 in reading to .51 for the total scale score. Most of the variance is explained by the baseline scores (either the first grade pre-test or the first grade post-test). "Family Income" , "Attendance", and "Race" show some relatively large effects, and these effects are mostly statistically significant. Most importantly, membership in SAGE schools has a consistent statistically significant positive effect on all sub-tests and the total score. A negative relationship suggests that, if these conditions are present, the test score will be lower.

Table 16. Scale Scores Regression – Second Grade Block Three Unstandardized Coefficients: First Grade Pre-Test as Control, 1999-00

	Reading	Language Arts	Math	Total
Pre-Test Score	.346*	.426*	.571*	.568*
Days Absent	-.411*	-.451*	-.553*	-.384*
Lunch Eligibility	-5.533*	-3.046*	-3.334*	-2.963*
African American	-5.179*	-13.376*	-8.475*	-9.343*
White	1.375	-1.452	2.459	-.899
SAGE	5.305*	9.187*	17.391*	10.417*
Constant	434.264*	388.481*	286.088*	302.887*
Adjusted R Squared	.233	.259	.391	.423
Standard Error	29.59	34.70	32.13	25.27
*Significant at .05 level				

Table 17. Scale Scores Regression – Second Grade Block Three Unstandardized Coefficients: First Grade Post Test as Control, 1999-00

	Reading	Language Arts	Math	Total
Pre-Test Score	.400*	.484*	.592*	.639*
Days Absent	-.439*	-.282	-.491*	-.298*
Lunch Eligibility	-3.714*	-3.128*	-3.218*	-1.837*
African American	-5.997*	-10.276*	-10.738*	-9.203*
White	2.934	.003	5.683*	1.606
SAGE	3.352*	6.987*	11.469*	6.029*
Constant	385.067*	337.194*	255.189*	237.600*
Adjusted R Squared	.286	.297	.448	.506
Standard Error	28.45	33.86	30.48	23.29
*Significant at .05 level				

Effect Sizes

Because the sample sizes used in these analyses were very large, even small group differences will result in statistically significant results. In order to better characterize the actual differences between groups, effect-size indicators were constructed. Effect sizes are interpreted as the group differences in terms of standard deviations. In general, effect sizes of .25 and below are considered modest, those from .25 to .50 are moderate, and those above .50 are large (Cohen, 1977). Two different indicators were used. First an "unadjusted" effect size was computed by dividing the difference between the SAGE and Comparison second grade test means by their pooled standard deviation. Because these means are affected by first-grade pre-test, SES,

race/ethnicity, and attendance differences, a second effect size measure was computed adjusting for these differences. This second measure used the raw score regression coefficient for the SAGE dummy variable in the regression analysis as an adjusted mean difference and divided this by the pooled standard deviation. The results of these computations for second grade are presented in Table 18 and should aid in the evaluation of the practical significance of the class-size advantage.

Table 18. 1999-00 Adjusted and Unadjusted Effect Sizes, Grade 2

Mathematics		**Reading**		**Language Arts**		**Total Score**	
Adjusted	**Unadjusted**	**Adjusted**	**Unadjusted**	**Adjusted**	**Unadjusted**	**Adjusted**	**Unadjusted**
.427	.308	.157	.063	.230	.171	.315	.213

African American Students

A precursor to the SAGE program is the Tennessee STAR experiment in reduced class size, a statewide initiative involving 7,000 students over four years in grades kindergarten through grade 3. One of the conclusions reached in the Tennessee experiment in reduced class size is that "the advantage of being in a small class is greater for minority students than for Whites" (Finn & Achilles, 1990, p.567). Note that no distinction is made among minority subgroups. In the SAGE evaluation, some of the racial/ethnic variables create complications. The variable "Asian" is a gross indicator that fails to distinguish among various Asian subgroups. For example, we are unable to distinguish Hmong students, who tend to be more disadvantaged, from other Asian subgroups. Native Americans are only minimally represented in Comparison schools. And many Hispanic students are limited in English proficiency and did not take the CTBS.

African American students comprise the largest minority subgroup of second grade test scores – roughly 23% of all SAGE students and 27% of all Comparison students. In the analyses to follow, African American students are first compared across SAGE and Comparison schools on the CTBS sub-tests and total scale score. Second, African American students are compared to white students across SAGE and Comparison schools on the CTBS Total Scale Score.

Table 19 provides comparisons of means on the CTBS second grade test, as well as change scores from the first grade pre-test to the second grade test and from the first grade post-test to the second grade test. On the second grade test, African American SAGE students scored significantly higher than African American Comparison school students on every sub-test and on the total scale score.

When using the first grade pre-test as the baseline score, statistically significant change scores are also found on all scores except for language arts. Using the first grade post-test as the baseline score shows that African American SAGE students made a significantly larger gain than their Comparison school counterparts in mathematics. SAGE students also made higher gains on the total scale score as well as language arts, but the gains are not statistically significant. In other words, African American SAGE students consistently outperform African American comparison school students with the significant gains being made during the first grade.

Table 19. African American Post-Test and Second Grade Change Scores, by SAGE and Comparison Schools, 1999-00

Score	SAGE	Comparison	Difference
Language Arts			
Mean Second Grade Score	602.90	589.15	13.75*

Mean Change From First Grade Pre-Test to Second Grade	65.45	60.58	4.87
Mean Change From First Grade Post-Test to Second Grade	20.83	19.95	0.88
Reading			
Mean Second Grade Score	602.32	595.11	7.20*
Mean Change From First Grade Pre-Test to Second Grade	70.92	59.60	11.32*
Mean Change From First Grade Post-Test to Second Grade	23.81	24.40	0.59
Mathematics			
Mean Second Grade Score	561.28	540.08	21.20*
Mean Change From First Grade Pre-Test to Second Grade	77.68	51.61	26.07*
Mean Change From First Grade Post-Test to Second Grade	28.53	18.66	9.87*
Total			
Mean Second Grade Score	589.42	575.33	14.09*
Mean Change From First Grade Pre-Test to Second Grade	71.21	57.80	13.40*
Mean Change From First Grade Post-Test to Second Grade	23.72	21.87	1.85
*significant at .05 level			

African American and White Achievement. African American students, as a group, scored lower than white students on the first grade pre-test total scale score, as shown in Table 20. This difference is statistically significant for both SAGE and Comparison schools, though

the gap between African American and white students is larger in the Comparison schools. The change from the first grade post-test to the second grade test shows that the SAGE African American students kept pace with white students but did not further close the achievement gap. African-American students in Comparison schools also kept pace with white students and did not fall further behind.

Table 20. African American versus White Achievement on Total Scale, 1999-00

	First Grade Pre-Test	**First Grade Post-Test**	**Second Grade**	**Change From Pre-Test to Second**	**Change From Post-Test to Second**
SAGE					
African American	516.83	564.21	589.42	71.21	23.72
White	538.87	580.34	610.18	71.09	29.48
Difference	22.04*	16.13*	20.76*	0.12	5.76*
Comparison					
African American	514.40	554.16	575.33	57.80	21.87
White	540.07	576.36	605.56	64.96	28.78
Difference	25.67*	22.20*	30.23*	7.16	6.91*
*significant at .05 level					

Third Grade Results 1999-00

Descriptive Statistics

Valid Test Scores. Analyses were conducted to assess the impact of SAGE on the 1999-00 third grade CTBS Complete Battery, Terra Nova Level 13 post-test results. The number of

third grade students with valid test scores for the Fall 1997 first grade pre-test, the Spring 1998 first grade post-test, the Spring 1999 second grade test, or the Spring 2000 third grade test are presented in Table 21.

Table 21. Number of 1999-00 Third Grade Students with Valid First Grade Pre-Test, First Grade Post-Test or Second Grade Test Scores

	Fall 1997 First Grade Pre-test		**Spring 1998 First Grade Post-test**		**Second Grade Test 1999**		**Third Grade Test 2000**	
	SAGE	**Comp**	SAGE	**Comp**	**SAGE**	**Comp**	**SAGE**	**Comp**
Reading	789	416	797	433	954	712	1262	921
Language Arts	789	416	797	433	954	712	1262	921
Mathematics	786	420	801	431	957	706	1279	907
Total	782	411	795	427	948	698	1256	898

In the analyses to follow, third grade test results are compared to the first grade pre-test, the first grade post-test, and the second grade test. Therefore, only those third-grade students who were present for both the first grade pre-test and post-test, as well as the second grade and third grade tests, were used in the 1999-00 third grade analysis. As would be expected, the number of third grade students having all four valid test scores was substantially less than the total number of students. Those students present in both the 1997-98 SAGE and Comparison first grade classrooms, the 1998-99 SAGE and Comparison second grade classrooms, and the 1999-00 SAGE and Comparison third grade classrooms were termed “three-year persisters”. There were 1092 three-year persisters, of which 705 were students in 30 SAGE schools and 387 were students in 16 Comparison schools. It should be noted that some persisters did not take all four CTBS tests. Table 22 shows the number of 1999-00 third grade persisters who took the test battery or one or more of the sub-tests at each of the three grade levels.

Table 22. Number of 1999-00 Third Grade Persisters with Valid First Grade Pre-Test, Post-Test, or Second Grade Test Scores in Total Scale and/or one or more sub-tests

	Fall 1997 First Grade Pre-test		**Spring 1998 First Grade Post-test**		**Second Grade Test 1999**		**Third Grade Test 2000**	
	SAGE	**Comp**	SAGE	**Comp**	**SAGE**	**Comp**	**SAGE**	**Comp**
Reading	703	378	702	384	701	383	700	382
Language Arts	703	378	702	384	701	383	700	382
Mathematics	699	384	704	382	702	380	703	377
Total	697	375	701	379	698	376	698	372

First Grade and Second Grade (Baseline) Results. First grade pre-test, first grade post-test, and second grade test scores served as a baseline. Table 23 provides descriptive statistics on the scale scores from the first grade pre-test, first grade post-test, and second grade test. Table 24 provides descriptive statistics for the third grade test.

Table 23. Descriptive Statistics Means (and Standard Deviations) on CTBS First Grade Pre-Test and Post-Test and Second Grade Test (SAGE and Comparison), 1999-00

	First Grade Pre-Test		**First Grade Post-Test**		**Second Grade Test**	
	SAGE	**Comp**	**SAGE**	**Comp**	**SAGE**	**Comp**
Reading	539.89	541.89	586.66	578.70	613.23	608.12
	(34.34)	(31.61)	(37.09)	(40.09)	(33.32)	(32.94)
Language Arts	541.61	540.81	593.50	586.99	617.57	604.90
	(39.13)	(39.85)	(42.30)	(42.06)	(38.08)	(35.00)
Math	502.64	504.65	543.91	539.20	578.86	560.85
	(37.15)	(36.05)	(38.73)	(37.72)	(38.96)	(38.29)
Total	528.33	529.09	574.74	568.20	603.37	591.25
	(31.20)	(29.96)	(33.38)	(33.76)	(32.03)	(30.60)

Table 24. SAGE and Comparison 1999-00 Descriptive Statistics CTBS Third Grade Test

	Scale Scores				**Normal Curve Equivalent**			
	SAGE		**Comparison**		**SAGE**		**Comparison**	
	Mean	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D**	**Mean**	**S.D.**
Reading	634.73	37.75	630.58	36.99	52.54	18.76	50.50	18.82
Language Arts	631.62	33.88	624.99	33.83	51.77	17.39	48.29	17.16
Mathematics	609.47	34.04	597.59	36.24	51.21	17.28	45.44	18.00
Total	625.45	31.38	618.15	31.68	52.26	17.48	48.29	17.47

Difference of Means Test. Results from the difference of means tests between SAGE and Comparison student scores from the 1999-00 third grade test are reported in Table 25. Third grade SAGE students scored significantly higher than Comparison students in all areas of the test except for reading. However, these gains have not been corrected for pre-existing differences between groups on factors such as prior achievement, attendance, race and SES.

Table 25. Differences of Means Test 1999-00 Third Grade Test

	SAGE	**Comparison**	**Difference**
Reading	634.73	630.58	4.15
Language Arts	631.62	624.99	6.62*
Mathematics	609.47	597.59	11.88*
Total	625.45	618.15	7.30*
*significant at .05 level			

The results from the difference of means tests between SAGE and Comparison student scale scores from the Fall 1997 first grade pre-test, Spring 1998 first grade post-test, Spring 1999 second grade test and Spring 2000 third grade test are reported in Table 26. These results reflect comparisons on an individual student level, the differences in gain scores between SAGE and Comparison students.

When the first grade pre-test is used as the baseline score, SAGE students made significantly higher gains than did Comparison students in all sub-tests and in the total score. The largest gain in SAGE student scores was on the mathematics sub-test. The smallest relative gain for SAGE students was on the language sub-scale, but even this gain was significant. When the first grade post-test is used as the baseline score, significant results continue to be found on

the mathematics sub-scale. These results suggest that the positive effects of SAGE on the mathematics sub-scale were maintained and significantly increased in second grade. When the second grade test scores are used for a baseline, Comparison students began closing the gap. However, as shown previously in Table 25, SAGE students still performed significantly higher than Comparison students on the third grade test in all scales except for reading. Again, this suggests that the positive effects of SAGE are most noticeable in the first grade.

Table 26. SAGE and Comparison Unadjusted Gain for Third Graders, 1999-00

	From First Grade Pre-Test to Third Grade Test			**From First Grade Post-Test to Third Grade Test**			**From Second Grade Test to Third Grade Test**		
Scale Score	**SAGE Gain**	**Comp Gain**	**Gain Diff.**	**SAGE Gain**	**Comp Gain**	**Gain Diff.**	**SAGE Gain**	**Comp Gain**	**Gain Diff.**
Reading	94.69	88.46	6.23*	48.08	51.83	-3.75	21.47	22.72	-1.23
Language	89.85	84.01	5.84*	38.20	37.88	0.31	14.16	20.28	-6.12*
Mathematics	106.94	92.29	14.65*	65.24	58.54	6.70*	30.63	35.83	-5.20*
Total	97.03	88.27	8.76*	50.47	49.47	1.00	22.17	26.44	-4.27*
*significant at .05 level									

Regression Analysis

As in the unadjusted gains reported for second graders, the gains depicted in Table 26 above do not reflect group differences related to socioeconomic status, ethnicity, attendance and prior knowledge. In order to correct for group differences related to these factors, regression analysis was employed. Essentially, regression analysis allows for a statistical adjustment that "equalizes" the groups on factors where pre-existing differences exist.

Regression Models. The effect of the SAGE program on student achievement for third graders was also tested through a series of ordinary least squares regression models for each sub-test and total scale score. Control variables were again entered into the models in blocks, with

the SAGE/Comparison student variable entered into the models last. In addition, three different regressions were done for each sub-test and total scale score. The first regression used the first-grade pre-test as a predictor variable (Table 27), the second regression used the first grade post-test as a predictor variable (Table 28), and the third regression used the second grade test as a predictor variable (Table 29).

The first block of control variables included student scores on the first grade pre-test or post-test, attendance, and eligibility for subsidized lunch as an indicator of family income. As with the second graders (discussed earlier), the second block of control variables added dummy variables for race/ethnicity. Finally, a dummy variable for SAGE or Comparison school student was added on the third block. This variable is coded 0 if a student is from a Comparison school and 1 if a student is from a SAGE school.

Regression Results. Results of the regression analyses are presented in Tables 27-29. When the first grade pre-test is used as the predictor variable, membership in SAGE emerges as a significant predictor of student achievement on the total score and for all sub-tests. The magnitude of the effect of SAGE on student achievement, as denoted by the unstandardized regression coefficient, varies depending on the CTBS sub-test.

The largest effects of SAGE are found when the first grade mathematics pre-test is used to predict the third grade test (12.410). When all cases are analyzed, the goodness-of-fit of the models (as denoted by the adjusted R square statistic), ranges from .29 in reading to .63 on the Total scale. Most of the variance, as was the case with the second graders, is explained by the baseline scores (either the first grade pre-test, first grade post-test, or second grade test). "Family Income", "Attendance", and "Race" show some relatively large effects, and these effects are usually statistically significant. This suggests that a student with a high absentee rate or from a

low socio-economic status will have lower test scores.

Table 27. Scale Scores Regression – Third Grade Block Three Unstandardized Coefficients: First Grade Pre-Test as Control, 1999-00

	Reading	**Language Arts**	**Math**	**Total**
Pre-Test Score	.470*	.387*	.469*	.629*
Days Absent	-.215	-.590*	-.640*	-.505*
Lunch Eligibility	-4.422*	-3.602*	-4.430*	-2.451*
African American	-3.650	-4.103	-6.454*	-3.475
White	10.778*	5.162*	2.711	4.210*
SAGE	5.954*	5.517*	12.410*	7.928*
Constant	380.117*	424.597*	373.149*	290.780*
Adjusted R Squared	.293	.307	.380	.490
Standard Error	31.46	28.05	27.68	22.58
*significant at .05 level				

Table 28. Scale Scores Regression – Third Grade Block Three Unstandardized Coefficients: First Grade Post-Test as Control, 1999-00

	Reading	**Language Arts**	**Math**	**Total**
Pre-Test Score	.430*	.359*	.475*	.573*
Days Absent	-.147	-.414*	-.406*	-.267
Lunch Eligibility	-4.649*	-4.173*	-3.213*	-2.792*
African American	-4.925	-2.809	-6.944*	-4.400*
White	8.708*	5.249*	4.271	3.662
SAGE	1.025	3.544	8.677*	3.207*
Constant	388.222*	422.844*	349.955*	298.996*
Adjusted R Squared	.308	.308	.403	.483
Standard Error	31.00	28.00	27.04	22.66
*Significant at .05 level				

Table 29. Scale Scores Regression – Third Grade Block Three Unstandardized Coefficients: Second Grade Test as Control, 1999-00

	Reading	Language Arts	Math	Total
Pre-Test Score	.748*	.492*	.568*	.748*
Days Absent	.020	-.539*	-.367*	-.269*
Lunch Eligibility	-1.566	-3.306*	-2.134*	-1.111
African American	-2.999	-.028	-1.774	-0.380
White	5.007*	4.993*	4.652*	3.142*
SAGE	0.949	-.377	1.920	-1.506
Constant	176.723*	334.462*	283.163*	177.467*
Adjusted R Squared	.512	.385	.501	.629
Standard Error	26.02	26.35	24.84	19.14
*significant at .05 level				

Effect Sizes

As in the second grade, the sample sizes used in the third grade analyses above were very large. Consequently, even small group differences will result in statistically significant results. In order to better characterize the actual differences between groups, effect-size indicators were also constructed for the third grade. Effect sizes are interpreted as the group differences in terms of standard deviations. In general, effect sizes of .25 and below are considered modest, those from .25 to .50 are moderate, and those above .50 are large (Cohen, 1977). Two different indicators were used. First an "unadjusted" effect size was computed by dividing the difference between the SAGE and comparison post-test means by their pooled standard deviation. Because these means are affected by pre-test, SES, and attendance differences, a second effect size measure was computed adjusting for these differences. This second measure used the raw score regression coefficient for the SAGE dummy variable in the regression analysis as an adjusted mean difference and divided this by the pooled standard deviation. The results of these computations for third grade are presented in Table 30.

Table 30. 1999-00 Adjusted and Unadjusted Effect Sizes, Grade 3

Mathematics		**Reading**		**Language Arts**		**Total Score**	
Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted
.357	.341	.159	.111	.163	.195	.252	.232

African American Students

As in the second grade classrooms, African American third grade students comprise the largest racial subgroup of test scores – roughly 23% of all SAGE students and 23% of all Comparison students. In the analyses to follow, African American students are first compared across SAGE and Comparison schools on the CTBS sub-tests and total scale score. Second, African American students are compared to white students across SAGE and Comparison schools on the CTBS Total Scale Score.

Table 31 provides comparisons of means on the CTBS third grade test, as well as change scores from the first grade pre-test, first grade post-test, and second grade test to the third grade test. On the third grade test, African American SAGE students tended to score higher than African American Comparison school students on every sub-test and on the total scale score. As can be seen, the differences on the mathematics and language sub-scales and the total scale scores are statistically significant.

When using the first grade pre-test as the baseline score, statistically significant change scores are found on all scores, with African American SAGE students outperforming African American Comparison students. Using the first grade post-test as the baseline score shows African American SAGE students continuing to make statistically significant gains on the math sub-test. However, using the second-grade test as the baseline score shows no further significant differences in achievement gain between SAGE and Comparison African American students. This suggests that African American SAGE students made significant gains with respect to their Comparison-school counterparts during the first grade.

Table 31. African American Third Grade Test and Change Scores, by SAGE or Comparison, 1999-00

Score	SAGE	Comparison	Difference
Language Arts			
Mean Third Grade Score	620.28	612.10	8.18*
Mean Change From First Grade Pre-Test to Third Grade	94.92	74.94	19.97*
Mean Change From First Grade Post-Test to Third Grade	41.45	39.24	2.21
Mean Change From Second Grade Test to Third Grade	18.92	20.50	1.58
Reading			
Mean Third Grade Score	619.77	613.89	5.58
Mean Change From First Grade Pre-Test to Third Grade	93.89	75.74	18.15*
Mean Change From First Grade Post-Test to Third Grade	43.18	50.40	7.22
Mean Change From Second Grade Test to Third Grade	18.21	20.67	2.46
Mathematics			
Mean Third Grade Score	593.98	575.26	18.72*
Mean Change From First Grade Pre-Test to Third Grade	117.33	83.28	34.05*
Mean Change From First Grade Post-Test to Third Grade	67.45	55.56	11.90*
Mean Change From Second Grade Test to Third Grade	38.55	36.35	2.20
Total			
Mean Third Grade Score	611.88	600.69	11.19*
Mean Change From First Grade Pre-Test to Third Grade	102.13	77.76	24.38*
Mean Change From First Grade Post-Test to Third Grade	50.46	48.80	1.67
Mean Change From Second Grade Test to Third Grade	25.70	26.28	0.58
*significant at .05 level			

African American and White Achievement. African American students, as a group, scored lower than white students on the total scale scores at each grade level, as shown in Table 32. This result is statistically significant for both SAGE and Comparison schools. African American students continued to score significantly lower than white students on total scale score and on all sub-tests, regardless of whether they were SAGE or Comparison school students. Gains made by African American versus white students were significantly better in SAGE schools from the beginning of first grade to the end of third grade. The opposite pattern was observed in Comparison schools.

Table 32. Third Grade African American and White Achievement on Total Scale, 1999-00

	First Grade Pre-Test	**First Grade Post-Test**	**Second Grade Test**	**Third Grade Test**	**Change From First Grade Pre-Test to Third Grade**	**Change From First Grade Post-test to Third Grade**	**Change From Second Grade to Third Grade**
SAGE							
African American	509.74	560.54	586.24	611.88	102.13	50.46	25.70
White	538.74	585.26	612.93	634.60	95.59	49.14	21.64
Difference	29.00*	24.72*	26.69*	22.73*	6.54*	1.32	4.07
Comparison							
African American	521.82	551.74	575.32	600.69	77.76	48.80	26.28
White	536.63	578.97	602.60	628.48	91.33	49.37	25.64
Difference	14.81*	27.23*	27.28*	27.80*	13.57*	0.57	0.64
*significant at .05 level							

Hierarchical Linear Modeling

Many social science research analyses involve hierarchical data structures. Hierarchical data structures are those in which individual units are nested within larger units, the latter being the unit of interest. The SAGE data are a prime example: students are nested within classrooms, and it is the classroom effect that is of particular interest to the SAGE project. Hierarchical data structures pose special analytical challenges in that data analysis at the individual level may result in a biased impression of the effect of the nesting unit (in the SAGE case, the classroom). An analytical approach known as "hierarchical linear modeling" or HLM (Bryk & Raudenbush, 1992) was specifically designed to accommodate these types of data structures. HLM was used with the SAGE data to provide an alternative and less biased account of the effects of SAGE experience on test scores. In these models, variables associated with individual students are referred to as level-1 variables and those associated with the classrooms are referred to as level-2 variables.

Analyses were conducted for each of the relevant criterion post-test scores: reading, mathematics, language arts, and total. For all analyses, the level-1 variables were pre-test achievement scores, SES measured as eligibility for subsidized lunch and "attendance" represented by days absent. In some cases this latter variable was omitted from the level-1 analysis due to multicollinearity with the SES variable in some classrooms (i.e., the attendance and SES variables were very highly correlated in some classrooms creating technical problems with the analysis). For second grade students, two sets of analyses were done, one using first grade pretest scores and the other first grade post test scores as the baseline measure. For third grade classrooms, three such analyses were done, one using first grade pretest, one using first grade post test and another using second grade post test as baseline achievement measures.

The post-test scores were adjusted for these three (two) level-1 variables at the individual level in each analysis, therefore the effects may be thought of as being statistically independent of the effects of these variables. Three different level-2 models are reported here: one specifying "class size" as the only level-2 variable, one including both "class size" and "SAGE" variables, and the last including both "class SES" and "SAGE" variables. Class SES was computed as the student SES average within each class and was therefore a measure of class poverty. Thus, for each subject area and grade, multiple analyses were done utilizing each of the baseline measures for each of the three "models" described below.

Tables 33-37 provide a summary of the effects of each of the level-1 and level-2 variables for each of these analyses. Level-1 effects can be interpreted as the weighted average of the within-classroom effects of the level-1 variables. Level-2 effects can interpreted as the classroom effects of the level-2 variables. The level-1 results indicate that within classrooms lower individual SES is related to lower post-test scores and higher pre-test scores are related to higher post-test scores. The coefficients associated with the level-2 variables can be thought of as classroom effects.

Model A. *Class Size*. These models examined the effect of class size on the adjusted criterion score. Class size equals the number of students divided by the number of teachers. The coefficient for the size variable can be interpreted as the loss (all coefficients were negative) in post test score attributable to the addition of an additional student to the classroom. For example in 2nd grade mathematics when first grade pretest was used as the baseline measure, each additional student added to the classroom could be expected to result in a .9521 decrease in the average mathematics post test score for the classroom. The results for all scores show this effect to be significant for all analyses except for the 2nd grade cohort in reading.

Model B. *Class Size, SAGE*. These models examined the effect of SAGE participation on the adjusted criterion score after the classrooms were class size adjusted, viewed as the effect of SAGE participation beyond the class size effect. Combining class size and SAGE participation in a single analysis isolates the effects that SAGE might have beyond those produced by lower class size. It should be noted that class size and SAGE are highly correlated variables and therefore the results are likely affected by multicollinearity problems (e.g., the coefficients are likely unreliable). The results show in general that once class size has been accounted for, SAGE has no significant effect on class average performance. This may suggest that the other SAGE interventions (i.e., rigorous curriculum, lighted school house, and staff development) are not having a significant impact on achievement in SAGE classrooms.

Model C. *Class SES, SAGE*. These models examined the effect of SAGE on the adjusted criterion score after the classrooms were SES adjusted, viewed as the effect of SAGE once the effects of the classroom SES are removed. Since socioeconomic status is known to have an influence on academic test scores, a replacement for this variable was used as both a level-1 and level-2 predictor. The level-2 variable was the average SES for the class and estimates the effect of the overall class SES level beyond that associated with the individual, which is accounted for in the level-1 model. This model combines class SES and SAGE. The results indicate that class SES has a significant effect on the class average post-test performance in all cases. The effect of a 1 point class average gain in SES equates to between a 13 point and 21 point gain on the average post-test score, depending on the sub-score. SES was measured on a three-point family income scale, thus a one point difference on average would be quite pronounced. The SAGE effect was significant in all cases except the third grade reading models once SES has been accounted for. Within each subtest, it is noteworthy that the SAGE effect in

this model remained relatively constant regardless of which baseline measure was used. This suggests that the classroom differences that exist due to SAGE are relatively constant from one grade to the next.

Table 33. HLM Results for 1999-00 Second Grade Students – First Grade Pre-Test as Initial Achievement

Source	**Total**	**Reading**	**Language Arts**	**Mathematics**
Level 1				
Pre-Test	.4868	.2836	.3756	.4422
SES	-2.7160	-6.0231	-3.9062	-5.1922
Attendance	-.5028	-.3512	**	-.9393
Level 2				
A. Class Size	-.4882*	-.1573	-.5717*	-.9521*
B. Class Size	6.7094	2.2373	7.7478	10.7154*
SAGE	-.1197	-.0285	-.1423	-.3440
C. Class SES	-15.7863*	-15.9516*	-16.6022*	-19.9922*
SAGE	10.6922*	5.4463*	12.1337*	16.8795*
*significant at .05 level **omitted due to multicollinearity				

Table 34. HLM Results for 1990-00 Second Grade Students – First Grade Post Test as Initial Achievement

Source	**Total**	**Reading**	**Language Arts**	**Mathematics**
Level 1				
Post-Test	.4438	.2929	.3363	.5052
SES	-5.0639	-6.7054	-5.2055	-4.7311
Attendance	-.6539	-.4114	**	-.9014
Level 2				
A. Class Size	-.7386*	-.2378	-.7834*	-1.2311*
B. Class Size	5.5510	2.6064	4.6951	9.2771
SAGE	-.4249	-.0869	.4025	-.6932
C. Class SES	-18.4729*	-16.9151*	-19.5780*	-21.8915*
SAGE	12.4198*	6.1849*	13.8869*	18.3259*
*significant at .05 level **omitted due to multicollinearity				

Table 35. HLM Results for 1999-00 Third Grade Students: First Grade Pre-Test as Initial Achievement

Source	Total	Reading	Language Arts	Mathematics
Level 1				
Pre-Test	.3247	.3003	.2974	.3194
SES	-6.1974	-6.4753	-6.4030	-7.6140
Attendance	-.2084	.5206	.2853	-.1843
Level 2				
B. Class Size	-.9909*	-.6439*	-.7048*	-1.6133*
B. Class Size	-3.3204	-3.2209	2.1889	-4.6652
SAGE	-1.2107*	-.8575	-.5569	-1.9346*
C. Class SES	-17.4872*	-18.4771*	-14.26281*	-18.3204*
SAGE	8.2123*	5.2296	7.4916*	12.7275*
*significant at .05 level				

Table 36. HLM Results for 1999-00 Third Grade Students – First Grade Post-Test as Initial Achievement

Source	Total	Reading	Language Arts	Mathematics
Level 1				
Pre-Test	.4123	.2980	.2666	.3972
SES	-5.0507	-7.0244	-6.3448	-5.1787
Attendance	.1655	**	-.03224	-.2044
Level 2				
A. Class Size	-1.0396*	-.7450*	-.7517*	-1.6552*
B. Class Size	-3.1204	-4.2077	2.4160	-4.8805
SAGE	-16.4241*	-1.0305*	-.5855	-1.9913*
C. Class SES	-16.4241*	-18.4269*	-13.9766*	-18.4237*
SAGE	8.0561*	5.5841	7.5554*	12.7143*
*significant at .05 level **omitted due to multicollinearity				

Table 37. HLM Results for 1999-00 Third Grade Students – Second Grade Test as Initial Achievement

Source	Total	Reading	Language Arts	Mathematics
Level 1				
Pre-Test	.5081	.5141	.3596	.4541
SES	-3.6032	-4.6556	-5.2532	-3.6290
Attendance	-.1908	.1922	-.2374	-.4638
Level 2				
A. Class Size	-.9946*	-.6567*	-.7067*	-1.6405*
B. Class Size	3.8605	-4.5368	2.1358	-5.1533
SAGE	-1.2534*	-.9274*	-.5594	-1.9948*
C. Class SES	-17.7009*	-19.3284*	-15.3101*	-18.9615*
Class Size	8.0417*	4.5429	7.4250*	12.6876*
*significant at .05 level				

Additional Analyses

The 1999-00 SAGE data base provided the opportunity to re-examine some specific factors which might be related to student performance within SAGE schools at grades two and three. Specifically, the following research questions were addressed:

1. Is the number of years of SAGE program participation related to individual or class achievement gains in grades two and three?
2. Is the socio-economic status (as measured by participation in the school lunch program) of SAGE participants related to individual achievement gains in grades two and three?
3. Is the type of SAGE classroom configuration related to classroom achievement gains in grades two and three?

A description of the analytical method and a summary of the results of these analyses is presented below for each question.

Years of Participation.

The relationship between the number of years of participation in the SAGE program and academic achievement gains was examined at both the individual student level and the aggregate classroom level. For both grades two and three, the scores from spring testing at each grade served as the variable of interest. Previous grade post-test, individual attendance, SES (as measured by school lunch participation), and number of years of SAGE participation were all used to predict the next grade's test score. This latter variable is the factor of interest. Table 38 shows the number of cases analyzed at each grade level by the number of years of SAGE participation.

Table 38. Number of SAGE Participants by Grade Level and Years of Participation, 1999-00

Years Participation	**1999-00 Grade Two**	**1999-00 Grade Three**
One	*	*
Two	134	123
Three	783	94
Four	--	600
Total	917	817

*Participants of one year's duration could not be analyzed due to lack of a pre-test

The results of this analysis showed no statistically significant relationships between years of program experience and achievement gain in any of the content areas when previous grade spring test, current grade spring test, attendance and SES were controlled with one exception. The Language Arts for third graders had a significant negative relationship with years in SAGE. Third grade Mathematics and Total scores also were negative, but they were not statistically significant. All results at grades two were in the expected positive direction, but none was statistically significant.

Analysis of years of participation in the SAGE program was also examined at the classroom level. For these analyses, variables were computed to represent average classroom levels for each of the variables used in the individual analysis. That is, previous grade spring test, current grade spring test, attendance, SES and years of SAGE experience were computed for each of the SAGE classrooms. Analyses similar to those done at the individual level were then done at the classroom level. For these analyses, there were 91 second grade classrooms and 83 third grade classrooms. In third grade classrooms, no significant relationships emerged between average years of SAGE participation and classroom achievement gains on any of the CTBS scores. For grade two, significant results were found for the reading sub-test and the total score (which can probably be attributed to the reading portion). Here a significant positive relationship existed indicating that those classrooms with a higher proportion of SAGE experienced students outperformed those classrooms with lower proportions of SAGE experienced students. Both the mathematics and language sub-tests also showed positive, although non-significant, relationships.

Socio-Economic Status.

The relationship between socio-economic status (as measured by the lunch participation variable) and academic achievement gains was examined at the individual student level at each grade. Regression analyses were done for each CTBS sub-test and the total score in order to address this question. Second grade and third grade scores served as the dependent variables. The independent variables included the previous grade post-test, individual attendance, and SES (as measured by school lunch participation). This latter variable is the factor of interest.

The results of these analyses are summarized in Table 39, which shows where significant relationships were found for each CTBS test and grade level. In all cases, the results show a negative relationship indicating that those with a lower SES index (higher actual SES) outperformed those with a higher SES index (lower actual SES).

Table 39. Significant Relationships for CTBS and SES by Grade Level, SAGE Students, 1999-00

CTBS Subtest	**99-00 Grade Two**	**99-00 Grade Three**
Reading	*	
Language	*	*
Mathematics	*	*
Total	*	
* significant negative relationship found		

Type of Classroom.

The implementation of the SAGE reduced class size feature has taken a number of forms. However, there are primarily two configurations: "true" 15:1 ratio classrooms where an individual teacher has 15 or fewer students and 30:2 ratio classrooms where two (or more) teachers have been given responsibility for more than 15 students. As in the past, it was of interest to determine if there are any achievement advantages associated with either of these basic configurations. These analyses were done at the classroom level with average post test performance serving as the dependent variable in each case. Independent variables included the appropriate average pretest score and a dichotomous variable indicating classroom type. There were 66 second grade and 60 third grade SAGE classroom with a 15:1 ratio, and there were 25 second grade and 23 third grade SAGE classrooms with a 30:2 ratio. Statistically significant results were found in three areas of Grade Two as indicated in Table 40. In these cases, the 15:1 ratio classrooms outperformed the 30:2 ratio classrooms.

Table 40. Statistically Significant Results between SAGE 15:1 and 30:2 Classrooms, 1999-00

CTBS Subtest	1999-00 Grade Two	1999-00 Grade Three
Reading	*	
Language		
Mathematics	*	
Total	*	
* significant relationship found favoring 15:1 over 30:2		

ANALYSES OF SAGE CLASSROOMS AND SCHOOLS 1999-2000

In this section, the effects of the SAGE program on teaching, curriculum, staff development, and lighted schoolhouse services are reported. Data regarding the effect of reduced size classes on teaching were obtained from a study of selected classrooms and teacher and principal questionnaires. The teacher and principal questionnaires also provided data concerning curriculum, staff development, and lighted schoolhouse services.

Previous Years

During 1996-97, 1997-98, and 1998-99 classroom events data were collected through classroom observations, teacher interviews, teacher logs, teacher questionnaires, and principal interviews. These instruments revealed that the dominant characteristic of teaching in reduced size classes is individualization. When classes become small, teachers provide for individual student needs through one-to-one tutoring, small group activities, and total class instruction where each child receives attention. This increased use of individualization is occasioned by less time spent on discipline and more time available for instruction, greater knowledge of individual students, and more enthusiasm for teaching. Individualization along with a slight increase in the use of hands-on activities results in more and deeper content coverage which in turn, it is speculated, brings about greater achievement. The type of individualization that seems to occur, however, is more process than substance. Teachers basically use direct instruction methods to accomplish established grade-level curriculum. In an effort to confirm and extend these findings, case studies consisting of more intensive classroom observations and teacher interviews were conducted in 1998-99. The case studies were conducted in three SAGE schools each representing a different type of SAGE classroom configuration: 15:1 Regular, 15:1 Shared Space, and 30:2

Team Taught. In each school a first-grade, second-grade, and third-grade classroom were studied in depth.

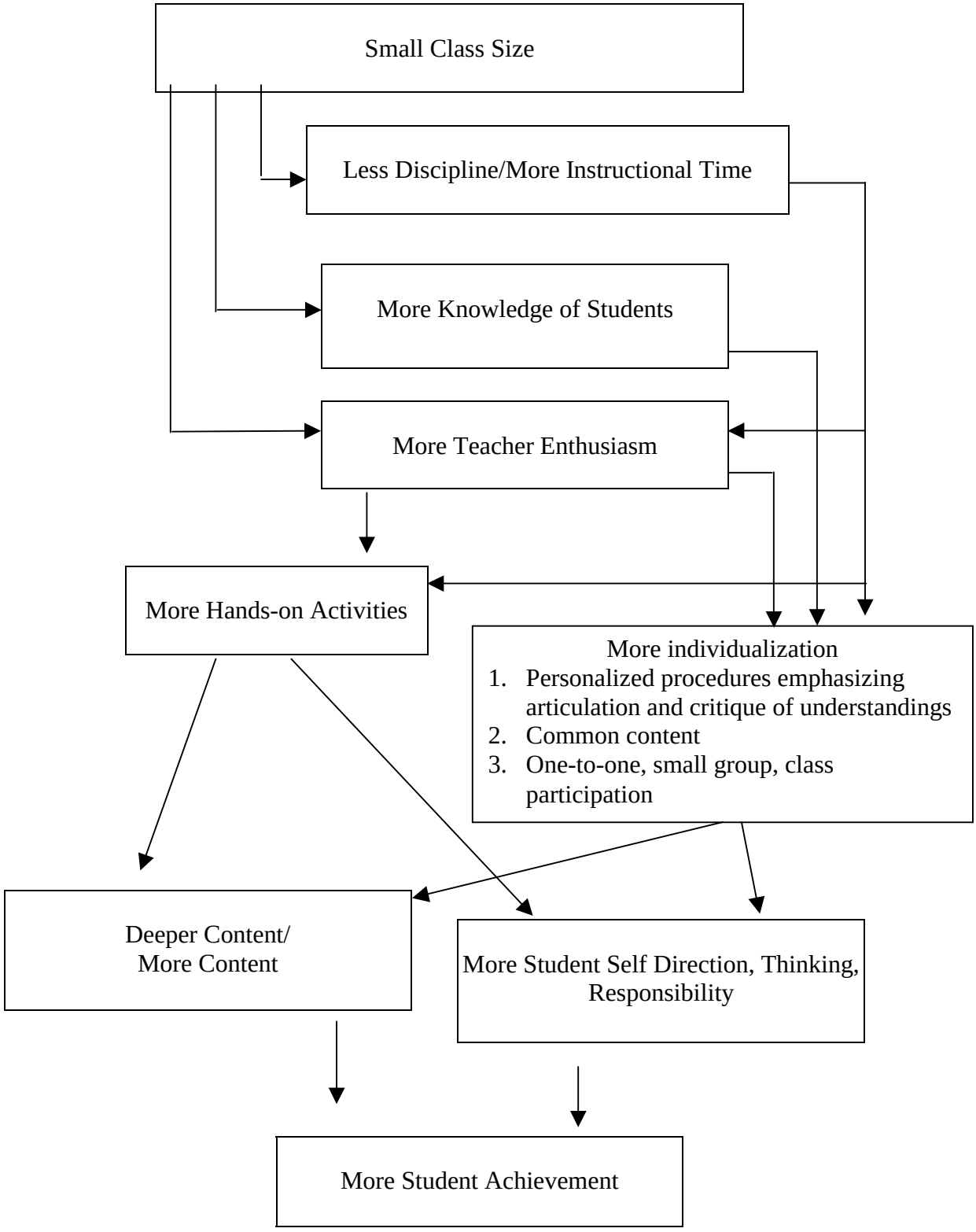
During 1999-00 the main focus of the classroom analysis component of the SAGE Evaluation Project was effective reduced class size teaching. Findings from the SAGE Evaluation Project indicate that student achievement is consistently higher in some SAGE classrooms than others. A sample of more effective and less effective SAGE first-grade classrooms was examined to identify teacher behaviors that are associated with higher student achievement. Knowing what effective reduced class size teachers do in their classrooms in comparison to less effective reduced class size teachers can provide the basis for reduced class size teaching staff development.

Rationale for the 1999-2000 Classroom Studies

Reducing class size to 15 students in first grade and in other primary grades improves student academic performance. The Student Achievement Guarantee in Education (SAGE) program has found with three cohorts of first-grade students totaling approximately 4500 students in 300 reduced size classrooms that achievement gains in reading, language arts, and mathematics were significantly higher than for a comparable group of regular class-size first-grade students. Scale score improvements for reduced class size students ranged from about 6 on an individual level of analysis to nearly 10 on a classroom level of analysis with SES and attendance controlled.

The SAGE evaluation also found that specific classroom events are associated with reduced size classes. The main effect of having fewer students is that teachers individualize their instruction, as can be seen in Figure 1.

Figure 1. A Model of Reduced Class Size Teaching and Learning



Teachers “tailor” their teaching to the needs of each student. Through one-to-one tutoring, small groups teaching and total class teaching, individual student’s understandings are elicited, critiqued, and corrected or extended. The content of instruction is uniform, but the teaching procedures vary with the student.

This increased use of individualization in reduced size classes is a result of increased knowledge of students, less discipline, which makes more time available for instruction, and greater teacher enthusiasm. The individualization that is produced, along with an increased use in hands-on activities that these three elements also enable, results in deeper and more content and in more student self-direction, and ultimately, we hypothesize, in greater student achievement as evidenced by higher achievement scores.

Although first-grade SAGE teachers out performed comparison schools teachers in terms of student academic achievement, variation in test score gains across SAGE first-grade teachers occurred. Some SAGE first-grade teachers were more effective than other first-grade teachers were. Instructional characteristics of SAGE teachers in general are known, but what more effective reduced class size teachers do in their classroom in comparison to less effective reduced class size teachers is not known. The purpose of this study was to identify teaching behaviors used by highly effective reduced class size first-grade teachers.

To determine the type of teaching used by highly effective reduced class size teachers, two analyses were carried out:

1. The teaching behaviors used by a group of highly effective, reduced class-size first-grade SAGE teachers were compared to the teaching behavior used by groups of less effective, reduced class size, first-grade SAGE teachers using qualitative research procedures.

2. The teaching behavior of all eligible reduced class size first-grade teachers as revealed by a teacher questionnaire were correlated with achievement test gain scores using quantitative research procedures.

Procedures

Subjects. The teachers for this study were 76 first-grade teachers or teacher teams who participated in the SAGE program for a minimum of two years. Using regression residuals of comparative levels of student achievement gains, teachers or teacher teams whose classrooms had comparatively higher than expected achievement gain scores for each of the two years and teachers or teacher teams whose classrooms had comparatively lower than expected achievement gain scores for each of the two years were identified.

Eight of the higher achievement teachers (the actual number of teacher was 10 because 2 classrooms were team-taught) and 5 of the lower achievement teachers were selected for qualitative study based on geographic accessibility. Two teachers, one from each group, were later dropped from the study because the teachers were not able to provide sufficient interview data for the study. The resulting two groups, then, consist of 7 higher-achieving teachers (5 are 15: 1 student-teacher ratio classrooms and 2 are 30:2 student- teacher ratio team taught classrooms for an actual total of 9 teachers) and 4 lower-achieving teachers. Two of the higher-achieving teachers are men. All of the lower-achieving teachers are women. The average years of teaching experience of the higher-achieving teachers is 14 and the average years of teaching experience of the lower-achieving teachers is 23. Two of the higher-achieving teachers have master's degrees and one of the lower-achieving teachers has a master's degree.

For the 11 teachers or teacher teams involved in the qualitative analysis, data were collected over a six-month period using classroom observations, teacher interviews, and a teacher self report. Of the total 76 first-grade SAGE teachers, 59 teachers or teacher teams completed teacher questionnaires in at least one of the two years. These 59 teachers or teacher teams were examined in the quantitative part of this study.

Classroom observation. Each teacher was observed a minimum of 4 times. Two of the observations were conducted in reading and 2 were conducted in mathematics. The observation guide (Appendix A) focused attention on general aspects of teaching as well as on teacher behaviors found to be related to reduced class size teaching from our previous research. During classroom observations, observers took notes which were used to prepare expanded descriptive accounts of the classroom events observed.

Teacher interviews. An introductory interview and one reading interview and one mathematics interview were conducted with each of the teachers. The introductory interview (Appendix B) obtained teacher background information, information about typical teaching characteristics, recent changes in teaching, information about class composition, and descriptions of normal testing procedures. The reading and mathematics interviews (Appendix C and D) asked teachers to describe their instructional philosophy regarding each of these areas, to describe a typical lesson in each of these areas, and to discuss perceived ways to improve their teaching in each of these areas. All interviews lasted from 30 minutes to over an hour and were tape-recorded and transcribed.

Teacher self-report. Along with the introductory interview, teachers were asked to complete a self-report regarding their instructional techniques and professional background

(Appendix E). This self- report focused on those teacher behaviors and their effects found to be representative of reduced class-size teaching.

The quantitative analysis involved the total group of 76 teachers. The SAGE database regarding achievement test scores, teacher questionnaire responses, and student profile reports was used to provide data.

Student achievement. Mean class total achievement gain scores were obtained for each of the 76 teachers for two successive years using the Comprehensive Test of Basic Skills (CTBS) Complete Battery, Terra Nova. Level 10 was administered in fall as a pretest and level 11 was administered in spring as a post-test. The mean gains for each of the two years were merged to arrive at a group mean achievement test gain score for each teacher.

Teacher questionnaire. A teacher questionnaire that had been administered to all SAGE teachers each spring has a section that asks teachers to rate and rank their use of 12 teaching behaviors that could, in theory, be expected to result from having a reduced size class (See Appendix). Teacher scores for each of the two years were averaged to provide a total teacher score for each item.

Student profile. The student profile is completed in fall and spring by each SAGE and Comparison school. It presents descriptive class information regarding size, enrollment, gender, race/ethnicity, SES, dominant language, and exceptionality. These data were again merged for the two-year period.

Qualitative Analysis Results

The general pattern of teaching found to be associated with teaching reduced size classes was evident in varying degrees in both the higher-achieving classrooms and in the lower-achieving classrooms. All of the teachers emphasized individualization to some degree. They all

attended to the needs of individual students through monitoring of learning, eliciting understandings, requiring students to display skills, providing feedback and critique, re-teaching when necessary, and in other ways. Discipline and management was less of an issue than it might have been in a larger class. Teachers expressed enthusiasm for teaching, although for some enthusiasm was less than in prior years in the SAGE program, probably because data regarding enthusiasm were collected in February rather than in May as in prior years. It was also found that teachers had knowledge of their students and employed hands-on activities, although not as often as they would prefer. The effects of deeper and more content coverage and growth in student self-direction and thinking were also found.

This pattern of teaching was not uniform across the two types of classrooms, however. Higher-achieving classrooms differed from lower-achieving classrooms in three dimensions: 1) instructional orientation, 2) management, and 3) individualization.

Instructional orientation.

Instructional orientation refers to ends and means preferred by the teacher. The ends are the goals or outcomes sought by the teachers. The means are the teaching methods or techniques used by the teacher to achieve the ends.

Lower-achieving teachers. Data revealed that the goals of lower-achieving teachers emphasized personal development. What these teachers wanted to accomplish with their students was to help students improve problem solving skill, develop critical thinking ability, gain deep understanding, enjoy learning, and in general become self-motivated, independent learners. Acquiring basic skills and fundamental concepts was not ignored, but it was secondary for these teachers. The teaching methods that they preferred were hands-on activities,

cooperative group work, problem solving tasks, and in general, child-centered experiential learning in which the teacher serves as a facilitator.

The instructional orientation of each of the lower-achieving teachers revealed one or more aspects of these kinds of goals and methods. Teacher L3 said, "*I try and be hands-on. I try and explain things in a way so that everybody has the opportunity and understands…So, I think the most dominant [characteristic] is my hands-on style."* She also stated, "*I like to get the kids problem solving. I like to not give them too much information. They're mainly having to come up with ideas themselves."* She commented further, "*I tend to not do skill and drill-type things, but to do small group problem solving type of problems."* L1's emphasis on problem solving and hands-on activities was evident in both her reading and mathematics instruction. Her room, which is organized into interest centers, provides many opportunities for students, individually and in small groups, to engage in experiential learning on their own. Further, the group activities she uses in mathematics often consist of solving problems, and the activities she uses in reading stress comprehension.

Both teachers L2 and L3 see fun and enjoyment as major goals in their teaching. Teacher L2 remarked "*One of the reasons I went into teaching is that I want kids to like reading. I'm not a reader. I don't like to read and that is one of the things for enjoyment… that students will read more if they enjoy it in the first place, so that is one of my main goals also."* She continued, "*I think that is a real important part of math that it should be a discovery part for them. It should be interesting, and it should be fun even though you - - you know, there are certain things, that you have to do…it is important to get the interest level."* Teacher L2 also said, "*I'm a real believer in having students feel comfortable and be able to express their feelings…I have a lot of hands-on things…Give them experience. That's one of the biggest things, I think."*

Teacher L3, although not neglectful of the need to memorize basic facts and acquire basic information, values creative teaching that produces student interest. She said, *"I'm creative. I don't know, you know how creative 'creative' is, but I'm sure that is one thing, you know, that they [colleagues] would say is pretty visible in my teaching…I can look past the typical type activities that you would do. For example, let's say in regard to reading a story, and do something using everyday materials that you wouldn't really think of but they just happen to be there, and it's something that would be fun and it sparks the interest of kids."*

Teacher L4 is dissimilar to the other teachers who had lower achievement in terms of having goals that emphasize thinking and other personal qualities or methods that emphasize hands-on activities and experiential learning. But, she is similar in that basic skills and concepts are not a primary goal for her. Teacher L4's main goal appears to be coverage. She said in regard to her teaching philosophy, "*Well, I try to get to the designated grade level at the end of the year, get through the book."* And, *"then they give us the [math] book and then this is what we have to take, we have to read and figure out what we're going to do. It's a lot of running off and a lot of preparation for everybody this year. Our math program, we have 6 of these books to go through."* This focus on coverage is coupled with an elaborate system for recording coverage progress and issuing grades. Recording student progress often substitutes for helping students see errors and re-teaching for Teacher L4.

Higher-achieving teachers. The instructional orientation for the higher-achieving teachers included personal goals and experimental methods but to a somewhat lesser degree than the lower-achieving teachers. The higher-achieving teachers have goals that are comprehensive but that emphasize basic skills and processes in general and in relation to the needs of individual students. Their reading programs focus on word attack skills and sight words along with

comprehension. Their mathematics programs are problem solving based, but they add an emphasis on basic facts and computational skills. This balanced set of goals is matched by a balanced set of instructional methods. The higher-achieving teachers have a variety of methods, including experiential methods; however, they give more attention to drill and practice than do lower-achieving teachers. And, when they do use discovery, problem solving, and other student centered techniques, they use them in a teacher-centered way. The teacher directs and controls the activity in such a way that predetermined objectives is reached.

This balanced instructional orientation with special attention given to the acquisition and practice of basic skills is descriptive of the teaching of each higher-achieving teacher. Regarding goals Teacher H3 said, *"I feel a reading program - - you need to have a balanced reading program where you have your phonics, you have some of your whole language… you need to have a lot of writing and reading, that they have a lot of practice."* Teacher H4 said concerning reading, *"So it's a combination of phonics, picture clues, and comprehension combined. I believe in it [phonics] very strongly… there's so many words that they are going to come across that if you were teaching completely sight vocabulary, they would have no skills to break down words."* She added, *"My strengths I think, are teaching basics and the structure."* Teacher H7 commented in regard to mathematics, *"I think it's great when we do it [hands on activities], but I also do the skills. I'm saying, they still have to know 2+2=4 in my room, and my students do know."* In regard to reading she said, *"I know people complain their series doesn't have a lot of skill work, but I give them skill work in addition on my own."* She concluded, *"So as you can see, I still do some skills, too, because I think it's important."* Remarks from teacher team H 6T also showed a basics orientation. They said, *"We do a lot of supplementing because we don't*

feel that the kids are all getting what they need from their series. In fact, they're missing a lot. A lot of stories don't reinforce the basic sight words and that's what we feel they need."

In regard to methods, most higher-achieving teachers' comments stress the need for both teacher-centered instruction and learner-centered instruction. Students need variety, but the variety must include drill and practice as well as explicit explanation and modeling. Teacher H2 said, *"In order to teach them how to read you're going to have to use a lot of different methods because children have different ways of learning."* Teacher team H5T said, *"You know you want to try to hit things from as many different angles as possible so that kids who need different angles pick up on things…we have to have a variety of ways to approach things."* Comments by the teacher team H6T and teacher H7 exemplify the higher-achieving teacher views about drill and practice. Teacher team H6T said their teaching is game oriented and they use lots of projects, but that, *" we still like the rote type activity because that's what math is all about."* Teacher H7 said in regard to reading, *" I could see at a glance who was catching on and who wasn't. And you know we'll go over it again and again."* In relation to explicit teaching, teacher H7 said, *"I don't know but I think it's better if they see a teacher model first. Oh, I really do. I know you're supposed to let them create and see what they come up with, but I think it just works better if the teacher models first and that's what I like to do."* Teacher H3 said, *" I do some direct instruction because I think they need that modeling. They need to be guided sometimes. They need that modeling and the redirection and checking for understanding."*

Management

Management, the second aspect of teaching in which higher and lower-achieving teachers differ, consists of student management and lesson management. Student management refers to classroom discipline polices and practices. Lesson management refers to the structuring, directing, and pacing of learning activities.

Lower-achieving teachers. The interviews and observations revealed that lower-achieving teachers differed from higher-achieving teachers in student management, lesson management, or both, in several ways. Lower-achieving teachers' student management can be characterized as being permissive and inconsistent. The lesson management of lower-achieving teachers revealed a tendency of the lessons to have unclear goals and to lack logical sequences. Teachers tend to pursue tangential ideas and to have lengthy exchanges with one or two students. The outcome of both of these sets of traits was student inattention and reduced engagement in learning tasks.

Of the four lower-achieving teachers, teachers L2 and L3 had difficulty in both student management and lesson management. Teacher L1's student management was effective, but her lesson management led to student confusion. And teacher L4's management, although unlike the management of the other lower-achieving teachers, had a similar effect on students.

Teacher L3's management style was evident in the opening day activities and a reading lesson that she taught. The day began as teacher L3 called the class of 14 students to the rug area. All of the students sat on the rug except for 2 who continued to walk around the main section of the room or into the reading "cubby hole" parts of the room that were formed by four-foot high bookcase dividers. The teacher did not require these 2 students to join the group. She generally ignored them, and possibly lost track of them, because once in a "cubby hole" area, they were not visible. The activities consisted of previewing the day's events by having students

place pictures and times on a chart at her direction (e.g., gym 12:30), discussing the date using a large calendar, and reading a story aloud to the class.

During the events and calendar activities many, if not all, of the students had opportunities to make contributions. Some students shouted out their comments, which were accepted by the teacher, while a few students waited for the teacher's recognition. The pace of both activities was slow and deliberate. She dwelled on topics with individual students for long periods of time while the rest of the class fidgeted, talked, and distracted each other. In determining for how many days school had been in session, for example, she carried on a dialogue with one student for 5 minutes at the front of the rug area. The other students appeared bored and many carried on side conversations. The discipline techniques teacher L3 used were positive and humane, (e.g. "James, when I hear you talking, I can't hear Michael."), but they were mostly ineffective in bringing about desired results. Not only did students talk among themselves; some left the group to wander about the room.

After the story had been read to the class and a brief discussion of it had taken place, the class was asked to return to their desks where the teacher and class engaged in another activity, this time about the weather. Again, many students chatted among themselves and some roamed about the room and generally ignored the teacher, or the teacher interacted with one student at length. As some students continued to misbehave, the teacher began to write students' names on the board, but compliance remained elusive.

The last activity was to write a story based on the story they had been read. The teacher modeled what she intended students to do, monitored the activity, and provided assistance when needed. In general students settled into the activity, although a few still walked around the room or had escaped to the "cubby hole."

The actual activities and methods that teacher L3 used were sound, and the stories that many of the students wrote were imaginative and detailed, but the effectiveness of the opening day activities was undoubtedly reduced by the teacher's acceptance and, in some cases, facilitation of student inattention. Her discipline procedures, lesson pacing, sequence of activities, and even room arrangement resulted in less time available for instruction.

Teacher L2's classroom management resembled that of teacher L3. She pursues individual students' comments that are only marginally related to the objectives of a lesson. She discusses them at great length while the class begins to unravel, she is inconsistent in enforcing her discipline policy of raising hands to seek permission to speak, and she dwells on topics beyond students' ability to attend.

The lesson management problems of teacher L1 consisted of vague goals for some of her lessons, a sequence of activities of tasks that lacked logic, and a slow pace. The slow pace and poor clarity of the lessons often resulted in interruption as students sought clarification and in inattention leading to incomplete or poorly completed tasks.

Teacher L4, as seen in relation to instructional orientation, has a teacher centered and teacher-controlled classroom. She has a discipline policy, which she follows consistently, and her lessons are carefully organized and sequenced. Her management, however, appears to be excessive. Lessons seem to progress regardless of student understanding. Misbehavior is dealt with instantly and sometimes harshly. She said to a student, for example, "*I don't want you to touch my stapler any more this year. Don't put your hand on my stapler again this year. If I had wanted you to staple them, I would have told you.*" In teacher L4's classroom, student learning and attention appear to be secondary to classroom efficiency and order. However, the punitive

and strict management style was ineffective in producing the desired results. The teacher appeared frustrated by classroom confusion and defiant students who refused to work.

Higher-achieving teachers. The higher-achieving teachers are considerably more able to manage students and manage lessons so that students are engaged in intended academic pursuits. Many see their ability to structure and organize as one of their most important teaching characteristics. Teacher H4 said, *"[Other people] would say I have a structured classroom. I try to maintain a routine because the children really need to have a routine. They need to know what to expect."* Teacher H1 remarked that what her colleagues would say about her was, *"Oh, she's structured, and I feel our kids, that's what they need, some structure. I try to keep a daily routine the same so that it isn't always changing, because the behaviors are such that they can't handle that."* Teacher H3 said, *"Another thing is I'm pretty organized, I have to be organized or it would drive me nuts."*

This ability of higher-achieving teachers to organize and manage was illustrated in a reading lesson taught by teacher H1. The lesson began with teacher H1 calling the class of 15 students together to listen to directions for the day's seatwork. As she waited for them to assemble, she reviewed the class rules for sharpening pencils. Also, in response to a child's request to get a drink, the teacher gently reminded him that students are not permitted to get drinks while the teacher is talking.

The seatwork consisted of four activities. The first, a several-part task involving vocabulary related to animals, was carefully explained and demonstrated by the teacher. After a part was explained step by step and the students begin that part, the teacher and classroom aide circulate and offer help where needed. The other three tasks were routine activities involving an

addition sheet, a word recognition sheet, and a story to be written in the students' journals. These tasks also were explained in detail.

Following the explanation of seatwork, interest centers to which students could go after completion of their seatwork were assigned and explained. The centers, which included a computer center, library-reading center, listening to a taped story center, and board activity center were assigned. Each student was told which center to attend. The centers were located in various sections of the room, but the children in the centers were visible from any area in the room. None of the bookcases and other dividers were tall enough to create hidden "cubby holes."

As the students begin the seat work the teachers calls the first of five reading groups to the front reading table. The groups, formed on the basis of reading ability, use different reading and instructional materials but follow a similar routine. Each includes vocabulary work; relating the story to students' experiences; predicting story events; oral reading either to the group, to the teacher, or in pairs; discussion of the story; and assigning of story-specific skill or comprehension exercises. During the reading group sessions the teacher continually surveyed the room and issued quick, decisive, but kindly, commands if students were becoming disruptive. She said, for example, *"Bruce, I shouldn't be able to hear your voice."* Also, as one group left the reading area and another group came to it, the teacher circulated around the room making sure each student was on task. When the last reading group returned to their desks, the teacher turned off the lights signaling the end of the reading period and time to put away or turn in their work.

Throughout this reading lesson all of the students are engaged and on task. The teacher has given clear directions, the tasks are appropriate and follow a logical progression, and the

pace is brisk. In her management of the students, teacher H1 is positive and nurturing, but she is also firm and decisive.

Although the management of all of the higher-achieving teachers results in a high degree of student engagement and production, teacher teams H5T and H6T have especially effective management. Student management is accomplished in a novel way in the teamed classrooms. One teacher is almost always available to oversee student attention and give help while the other presents a lesson. Further, because teamed teachers share their views of individual children, they are able to develop an in-depth understanding of each child and target responses to student inattention. In addition, however, each team has an elaborate student management system. Teachers H5T have developed a ticket-sticker system in which tickets are placed in a student's envelope for good behavior and removed for poor behavior. One of the teachers said, "*Three, two, one, we're done. Everyone freeze. I'll take a ticket from anyone not quiet."* But implementation of the system is neither rigid nor harsh. At another time he said, *"Now direct your attention up here. Please. Pretty please. Pretty please with sugar on top. Cherries?"* The other team's approach is based on self-control. The following comment was characteristic of their classroom; *"Someone in back is being disrespectful. Being disrespectful is making poor choices. It is important to always - - -," and* the class chimes in, *"make good choices."*

The teacher collaboration that produced these systems and their implementation also influences lesson management. Because lessons are planned together and are discussed and critiqued at length, they are usually well organized and efficiently presented. Unproductive tangents, ambiguous deviations, and slow pace rarely occur.

Individualization

The effect of an instructional orientation that emphasizes academic development and a type of management that enables it to flourish is, in the reduced size classroom, increased use of individualization. Individualization refers to meeting the needs of individual students by providing opportunities for them to reveal their understandings and abilities and offering critique and assistance in all settings. It occurs in both lower-achieving classrooms and higher-achieving classrooms almost automatically as a result of having a reduced size class. But, it occurs more often in higher-achieving classrooms than in lower-achieving classrooms.

Higher-achieving teachers. The higher-achieving teachers mention, and their teaching reflects, a high degree of individualization. They diagnose present levels of achievement, elicit students' thoughts, offer feedback, reteach when necessary, and give periodic reviews. Their lessons are characterized by a variety of types of activities in an attempt to facilitate various learning styles, by much sharing and oral reading, and by monitored practice. Teacher H2 said, *"You're going to have to use a lot of different methods because the children have different ways of learning."* She also said, *"I do some individual reading with every child. With this small class I can get around and listen to every child read individually every day."* Teacher H3 has regular writing conferences with her students individually which involve students reading their work to the teacher, students editing with the teacher, the teacher questioning students about the finished story, students sharing their interpretations, and the reading of the finished story to the teacher and, eventually, the class. She said her goal is *"that each child grows throughout the year… I just need to meet their needs."* Teacher H4 said she conferences with students individually about their books. She said she has the students *"reread [their books] to me. We talk about it. I can even question them about story maps. They read their favorite parts back.*

So I'm really doing a lot more one-on-one." Teacher H5T said in regard to variety in methods, *"You want to hit things from as many different angles as possible so that the kids who need different angles pick up on things. If you're forced to throw it down their throats in one way, you know, it would be like eating pudding all of the time. You know some kids can't eat pudding…we need to have a variety of ways to approach things."*

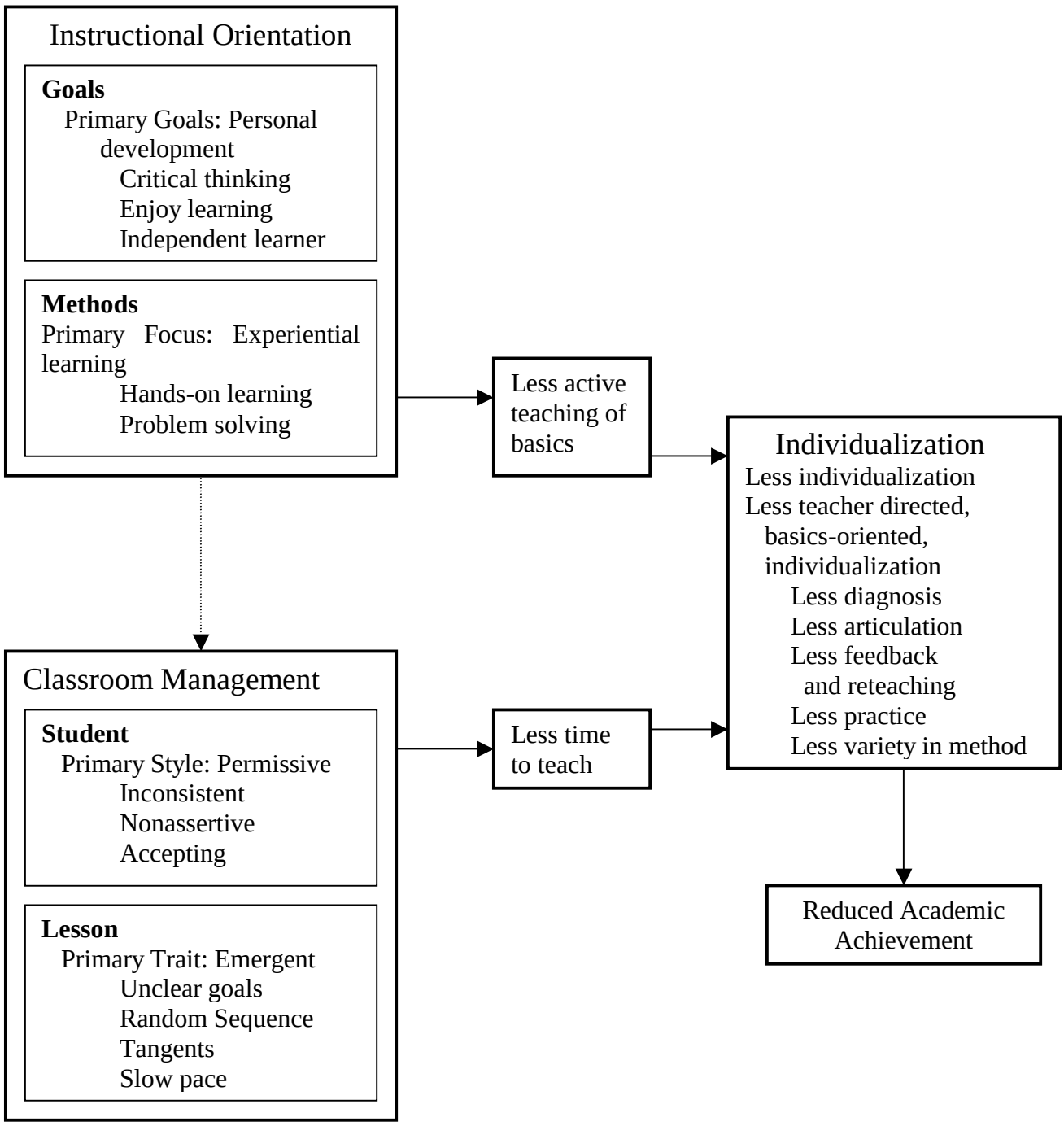
Lower-achieving teachers. The lower-achieving teachers are not greatly less inclined to focus on individuals, but their focus is less teacher directed and they have less time available to attend to teaching in general and to individuals in particular because of their management problems. Teacher L3, for example, uses a variety of methods, elicits students' understandings, has one-to-one sessions with students, and in other ways shows concern for individuals, but her individualization is not productive, and in some cases it is counterproductive. Her opposition to drill and practice and her inability to control the class negate any benefits of her individualization. And, as we have seen, when she does seek students' understanding or offers help to an individual student, she does it for an extended period of time, which causes misbehavior on the part of other students. This same problem affects teacher L1's teaching because her explanations are often confusing and L2's teaching because of her concern for students emotional well being. In one instance teacher L2, in discussing the story during a reading lesson, asked the class what they thought an enemy of the city was. One student said it was the car. Another, however, remarked that he did not like birds and began to tell the class why he did not like birds. The teacher encouraged the student to share his thoughts about birds at great length even though it was unrelated to the story and caused the class to become disruptive.

Tentative Models of Less Effective and More Effective Teaching

The tentative models depicted in Figures 2 and 3 represent two types of teaching that differ in degree. Both higher-achieving teachers and lower-achieving teachers use all of the elements identified in both figures, but the higher-achieving teachers use the elements in Figure 3 more often than the lower-achieving teachers and the lower-achieving teachers use the elements in Figure 2 more often than the higher-achieving teachers. Further, Figure 2 does not represent all of the lower-achieving teachers from whom data were collected. Teacher L4, as we have seen, is an outlier with different teaching characteristics. While she strove to teach the basics, her ineffective management style frustrated that effort; and the effects of her teaching match those of the other lower-achieving teachers.

Lower-achieving teachers, as can be seen in Figure 2, have goals that emphasize students' personal development and stress methods that facilitate independent, experiential learning. These preferences result in a less central role for the teacher and less emphasis on the basic skills and concepts of reading and mathematics in comparison to higher-achieving teachers. Also, lower-achieving teachers have student management procedures that are tolerant and permissive and lesson management practices that evolve and develop. These practices are time consuming and result in less time available to devote to academic, goal directed instruction in comparison to higher-achieving teachers. Although lower-achieving teachers use individualization in their reduced size classes, because of their attitude toward the active teaching of basics and the limited time available for instruction in their classrooms, their individualization is less teacher directed and basics oriented than higher-achieving teachers. The result, it is hypothesized, is reduced academic achievement.

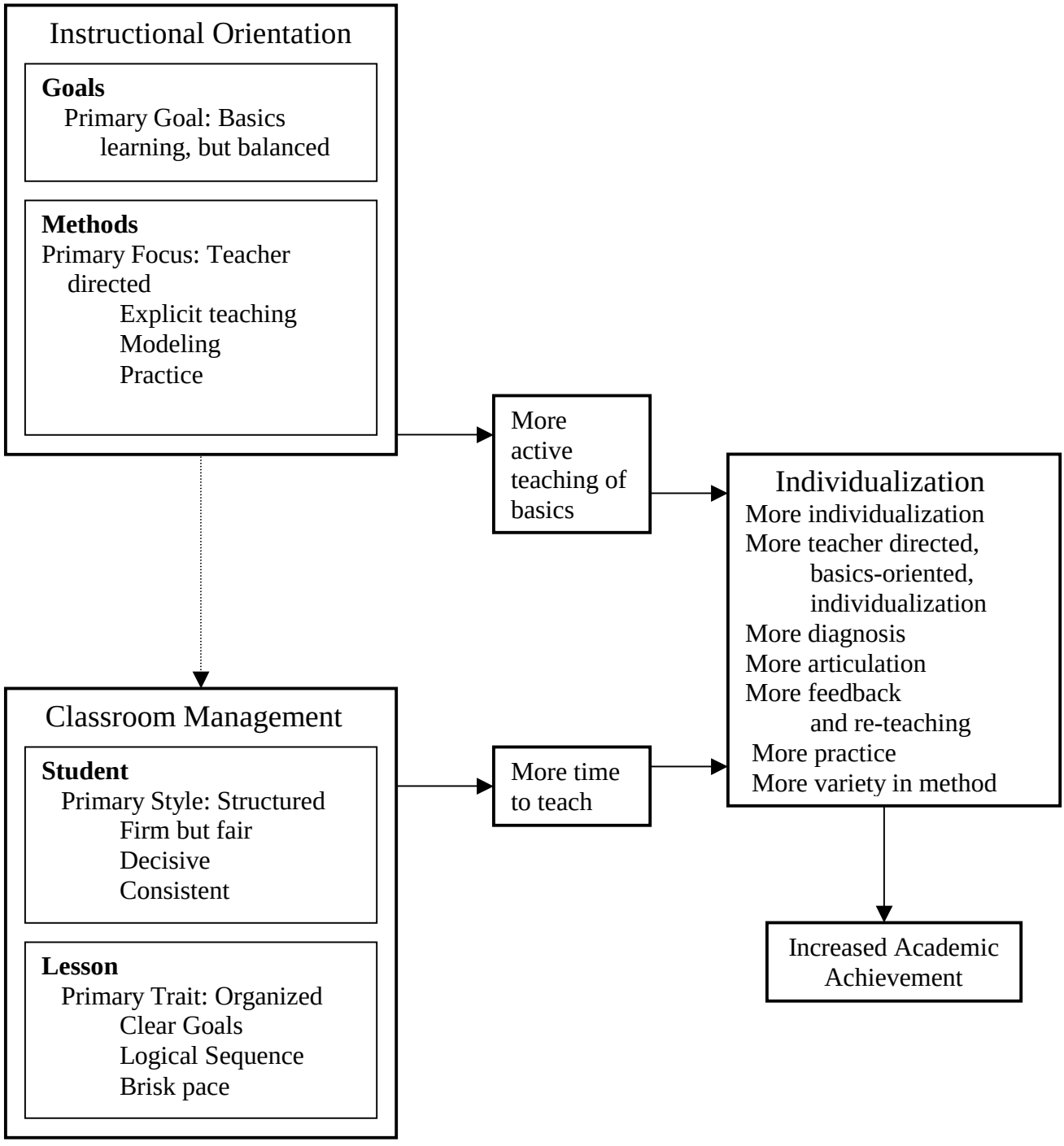
Figure 2. A Model of Less Effective Reduced Class Size Teaching



Higher-achieving teachers, as can be seen in Figure 3, have more balanced goals that include attention to personal development, but they emphasize the goals of basic skills and concepts. The methods that they prefer are those associated with explicit teaching such as explaining, modeling, checking, and evaluating. These goals and methods result in more active teaching of the basics in comparison to lower-achieving teachers. Concerning classroom management, higher-achieving teachers are structured and organized. Students are treated in a positive but consistent, firm way. Lessons are aimed at important goals and proceed in a systematic, efficient way. Together these student management techniques increase academic learning time as they decrease distractions. Because of the inclination to focus on the active teaching of basics and the increased amount of time available for instruction, the individualization of higher-achieving teachers is more teacher oriented and basics oriented than lower-achieving teachers. The result, again hypothesized, is increased student achievement.

It should be noted that although the teaching of lower-achieving teachers jeopardizes achievement as measured by standardized tests, over time the goals and methods of the lower-achieving teachers may not be harmful, and may, indeed, be helpful. If the goals of thinking and problem solving are realized, students will be served in the future even though the attainment of basics is delayed.

Figure 3. A Model of More Effective Reduced Class Size Teaching



Rival Hypotheses

The assumption on which this analysis of higher and lower-achieving reduced size classrooms is based is that teacher behavior in the two types of classrooms is the critical factor. Other factors could be responsible in whole or in part for the achievement differences, however. Three of the most plausible factors are students, aides, and testing. None of these factors was found to be a major cause of achievement differences.

Student ability and behavior could cause achievement differences between the two sets of classrooms, but differences between the two sets of classrooms on these variables did not exist. Since expected achievement used to identify higher and lower achievement teachers was based on pretest rather than post test achievement data, student ability differences in reading and mathematics were controlled. In terms of student behavior, in each of the 11 schools a stratified random procedure based on student behavior was used by the school to form classes. Students likely to misbehave and disrupt the class were equally distributed across classes. This procedure served to reduce or eliminate of behavioral problem differences across classes and between the two sets of classes.

A related student factor that could possibly explain achievement differences is number of exceptional education students in a classroom. Although exceptional education students were not used to calculate academic progress, in large numbers they could change the classroom dynamic. Here, too, however, students classified or likely to be classified as having exceptional education needs were usually distributed across classrooms rather than placed in one classroom. An exception was teacher L3's classroom. This teacher had special education certification and was

assigned a greater number of exceptional education students than other first-grade teachers in her school.

A classroom with many aides or assistants could be thought to have an advantage over a classroom where the teacher is the only adult providing instruction and help to students. The range of assistance varied from 16 hours a week to no hours per week in the 11 classrooms. Between the two sets of classrooms differences in amount of aide time were slight, however. The roles and responsibilities of the aides could also have an impact on achievement, but these data were beyond the scope of this analysis.

Substantive and procedural preparation for the achievement test could obviously affect the test scores. Every teacher in both sets of classrooms reported that she or he prepared students for the test by practicing filling in circles and other format features of the test and by trying to relax the students and make them emotionally ready to do their best. All of the teachers, except one, said that they did not specifically identify and teach content that would be tested. The exception, teacher H1, said, *"I'm always thinking back to the testing when I am teaching. What things need to be stressed? What do they need to do to learn…I've always looked at the state Terra Nova, and am I hitting all those points?"*

Teacher Behavior and Student Achievement Analysis

The results of the quantitative analysis of the relationship of teacher behavior to student achievement are reported in Table 41.

Table 41. Correlations of Ranking and Rating of Teacher Behaviors with Mean Classroom Achievement in Reduced Size Classes at the First-Grade Level.

	Ratings[2]	**Rankings**[3]
1. Spent more time teaching rather than managing the classroom	.220	.218
2. Covered more content	.017	-.195
3. Integrated content form several subjects	.162	.159
4. Covered content in more depth	.089	-.289*
5. Spent more time individualizing instruction, assessing learning, providing learning activities, and giving help.	.164	.172
6. Spent more time engaging students in discussion, encouraging them to share their ideas, and answering their questions	.101	.065
7. Involved students in more hands-on activities	.251	-.201
8. Based activities on students' prior knowledge, understandings, and skills	.271*	.096
9. More often involved students in problem solving, creating and experimenting	.207	-.036
10. More often organized the class into cooperative groups.	.126	.033
11. Offered more opportunities to choose among learning activities and materials	.076	-.112
12. Am more enthusiastic about my teaching1	.159	-.185

As can been seen, the correlations between the ratings of each of the 12 teacher questionnaires items and class achievement are positive. The rankings, where teachers were

Note:
[2] Ratings of teaching behaviors were obtained using a 5-point Likert scale.

[3] Rankings consist of the 3 most important teaching behaviors teachers used
* $\underline{P} < .05$

required to discriminate among the 12 items by identifying the 3 most important behaviors, reveal a different result. Six of the teacher behaviors correlate negatively with student achievement. As displayed in Table 41, these negative correlates are 1) covered more content; 2) covered more content in depth; 3) involved students in more hands-on activities; 4) more often involved students in problem solving, creating, and experimenting; 5) offered more opportunities to choose among learning activities and materials; and 6) more enthusiastic about teaching. The remaining six teacher behaviors that correlate positively with student achievement are 1) spent more time teaching than managing the classroom; 2) integrated subjects; 3) spent more time individualizing; 4) spent more time in discussion; 5) based activities on students' prior knowledge; 6) and more often used cooperative groups.

These two sets of correlates are generally consistent with the findings from the qualitative analysis regarding teaching behavior and serve to support and confirm those findings, particularly in regard to negative correlates and instructional orientation and to amount of individualization. Since the teacher questionnaire did not contain items related to specific forms of student classroom management, characteristics of lesson classroom management, and elements of teacher-directed methods, comparisons between qualitative and quantitative data cannot be made in these areas. Concerning methods, however, the opposite of hands-on activities, problem solving, and opportunity to choose is likely to be presenting, modeling, checking, and similar behaviors which would be consistent with the behaviors used by the more effective reduced class size teachers examined in the qualitative analysis.

Discussion and Implications

These findings, although tentative because of the limited sample size and the examination of only one grade level, have possible implications for staff development in reduced class size schools.

Improved teaching and learning seem to occur in most first-grade classrooms when class size is reduced to about 15 students. It is not unreasonable to speculate that even the teaching of less effective teachers improves as they move from a larger class to a smaller class. The results of this study suggest that these teachers as well as all reduced class size teachers could have a greater impact on students' learning if they employed particular kinds of instructional and management methods.

Just because having fewer students results in fewer student management problems and teachers can give students more freedom to explore on their own, to inquire independently, or to create without boundaries, it may not be in the students' best interests to do so. Having a small class is not a time for teachers to sit back and relax. It is not a time to be less assertive, less preplanned, and less focused because a small class permits these behaviors to be used without the danger of the class becoming out-of-control. Experiential learning and students' interest are, of course, important and need to be present in every classroom, but reduced class size teachers need to be cognizant of the unusual opportunity they have been given to advance the achievement of the individual students in their classrooms. They need to be encouraged to increase their emphasis on academic learning - not decrease it. They need to use more teacher directed, basics-oriented individualization with special emphasis on student articulation of understandings and teacher critique and re-teaching, not less.

Staff development programs emphasizing the teacher behaviors used by the more effective reduced class size teachers can conceivably strengthen the positive results that have been found to be associated with reduced class size. Making classes smaller is the first step. Helping teachers to improve their teaching is the second step.

Although the findings regarding effective reduced class size teaching are preliminary, they suggest that staff development programs in each SAGE school should focus its efforts to help reduced class size teachers use wisely the additional instructional time that reduced class size creates. The program needs to encourage teachers to redouble their efforts to complement a well organized and structured classroom where individuals are helped to acquire basic knowledge and skills.

Teacher and Principal Questionnaires

Self reported data from of all SAGE teachers regarding their teaching are contained in Tables 42-45. These data were obtained from the Teacher Questionnaire administered in the spring. Table 42 shows that for the total group of SAGE teachers, the teacher behaviors that received the highest ratings are *individualization* and *teacher enthusiasm* followed by *engaging students in discussion, using hands-on activities, teaching rather than disciplining*, and *covering more content*. Those behaviors receiving comparatively lower ratings are the more student-centered behaviors of *using cooperative groups* and *giving students choices in learning activities*. *Integrating content from several subjects, more time for teaching, problem solving activities, students prior knowledge, covering more content,* and *covering content in more depth* have ratings between the high and low groups of behaviors; however, 77% or more of the ratings for these items fall in the high group.

Table 42. Total Teacher Questionnaire Results, Grades K-3 (Percentages) 1999-00

ITEM	Strongly Disagree	Disagree	Sometimes	Agree	Strongly Agree	No* Response
1. More time teaching	0.5	1.5	12.4	40.5	35.9	9.3
2. Covered more content	0.2	2.2	10.7	42.4	35.1	9.3
3. Integrated content	0	0.5	13.2	44.6	32.2	9.5
4. More depth	0	1.2	10.2	46.6	32.2	9.5
5. Individualization	0.2	1.5	3.9	34.6	50.2	9.5
6. More engaging	0.2	0.7	4.9	38.5	46.3	9.3
7. More Hands-on	0	0.5	8.8	41.0	40.5	9.3
8. Student's knowledge	0	1.0	11.2	50.7	27.6	9.5
9. Problem solving	0.2	0.7	12.0	45.6	32.2	9.3
10. Cooperative groups	0.2	2.4	18.8	40.0	29.0	9.3
11. More opportunities	0	2.9	20.5	39.5	27.1	10.0
12. Teacher enthusiasm	0.5	1.7	5.9	31.5	51.0	9.5

N=410

*Teachers whose teaching experience has always been with a small class were asked to skip part A of the questionnaire since they lacked a comparison base needed to answer the questions for this part.

Teachers were also asked to select the three most significant ways their teaching has been affected by a reduced student-teacher rate. The rankings for the twelve items are shown in table 43. Teacher rankings of the most significant teaching behaviors related to smaller class sizes are similar to responses reported in table 42 with *individualization, more time for teaching, more engaging,* and *more hands-on* receiving high rankings.

Table 43. Total Teacher Questionnaire Results Rankings of 12 Items, Grades K-3 (Percentages) 1999-00

ITEM	Ranking*
1. More time teaching	15.4
2. Covered more content	6.9
3. Integrated content	3.9
4. More depth	5.8
5. Individualization	22.9
6. More engaging	10.1
7. More Hands-on	10.7
8. Student's knowledge	3.4
9. Problem solving	5.7
10. Cooperative groups	4.1
11. More opportunities	3.2
12. Teacher enthusiasm	7.6

N=410

*Average ranking of top three choices from the list of 12 items.

Teacher behavior by grade levels for first-, second-, and third-grade classrooms is reported in Table 44, and teacher behavior by type of SAGE classroom for these grade levels is reported in Table 45. It can be seen that the general pattern of ratings in the low and high groups of teaching behaviors as shown in Table 42 is also descriptive of each grade level and each type of classroom. However, as indicated in Table 44, some trends revealed by teacher responses of different grade levels are noteworthy. Findings by grade show that third-grade teachers appear less likely than first- and second-grade teachers to allocate ratings of strong agreement for the items of *covered more content, integrated content, individualization, more opportunities,* and *teacher enthusiasm.*

Table 44. Teacher Questionnaire Results for First, Second, and Third Grade (Percentages) 1999-00

	First Grade (N=79)					**Second Grade (N=89)**					**Third Grade (N=91)**				
	SD	**D**	**S**	**A**	**SA**	**SD**	**D**	**S**	**A**	**SA**	**SD**	**D**	**S**	**A**	**SA**
1	1.3	0	20.3	44.3	34.2	0	1.1	12.4	50.6	36	0	0	14.3	50.5	35.2
2	0	1.3	10.1	53.2	35.4	0	3.4	11.2	44.9	40.4	0	3.3	16.5	56.0	24.2
3	0	0	11.4	55.7	32.9	0	0	14.6	49.4	34.8	0	1.1	26.4	48.4	24.2
4	0	0	11.4	54.4	34.2	0	0	11.2	53.9	33.7	0	1.1	15.4	59.3	24.2
5	0	1.3	3.8	39.2	55.7	0	0	2.2	41.6	56.2	0	1.1	6.6	50.5	41.8
6	0	0	3.8	46.8	49.4	0	0	3.4	51.7	44.9	0	1.1	8.8	40.7	49.5
7	0	0	5.1	48.1	46.8	0	0	10.1	48.3	41.6	0	0	16.5	48.4	35.2
8	0	1.3	11.4	58.2	27.8	0	0	16.9	61.8	21.3	0	2.2	12.1	60.4	25.3
9	0	1.3	8.9	63.3	26.6	0	0	16.9	48.3	34.8	0	1.1	13.2	56.0	29.7
10	0	2.5	20.3	53.2	24.1	0	2.2	22.5	43.8	31.5	1.1	2.2	20.9	48.4	27.5
11	0	2.5	16.5	51.9	29.1	0	5.6	22.5	50.6	21.3	0	2.2	40.7	42.9	11.0
12	0	0	3.8	43.0	53.2	0	0	6.7	33.7	58.4	0	3.3	8.8	42.9	45.1

Key

SD = Strongly Disagree D = Disagree S = Sometimes A = Agree SA = Strongly Agree

This year five types of SAGE classroom configurations were reported by schools for first, second, and third grade classrooms: 15:1 regular reduced size; 15:1 shared space; 30:2 team taught; 30:2 classes with a floating teacher for reading, language arts, and math instruction, and classes with a full time and a part time teacher. The predominant type of classroom organization was 15:1 (144) followed by team taught classes (56). In a few classrooms, different types of

organization were used such as shared space (14), a floating teacher (4), and a full time/part time organization (7). As indicated in Table 45, the general pattern of ratings in the high and low groups of teaching behaviors also holds true for the findings by type of SAGE classroom. In Table 45, the results for floating teacher and full/part time teacher are combined as other types because both types reduce class size in similar ways.

Table 45. Teacher Questionnaire Results for Different Types of SAGE Classrooms (Percentages) in Grades 1-3 in 1999-00

	Regular 15:1 N=144					**Team Taught 30:2 N=56**					**Shared Space 15:1 N=14**					**Other Types N=11**				
	SD	**D**	**S**	**A**	**SA**	**SD**	**D**	**S**	**A**	**SA**	**SD**	**D**	**S**	**A**	**SA**	**SD**	**D**	**S**	**A**	**SA**
1	1	0	19	44	36	0	0	13	54	34	0	0	7	50	43	0	0	18	27	55
2	0	2	14	50	34	0	4	9	54	34	0	0	14	36	50	0	9	9	73	9
3	0	1	18	48	31	0	0	11	54	36	0	0	7	43	59	0	0	27	18	55
4	0	0	13	52	35	0	0	13	61	27	0	0	7	59	43	0	0	27	45	27
5	0	0	6	38	57	0	2	0	50	48	0	7	0	50	43	0	0	9	64	27
6	0	0	6	35	59	0	0	4	57	39	0	0	0	64	36	0	0	0	82	18
7	0	0	9	47	44	0	0	9	48	43	0	0	0	50	50	0	0	27	45	27
8	0	0	14	56	30	0	2	9	61	29	0	0	14	64	21	0	0	9	73	18
9	0	1	13	54	33	0	0	4	63	34	0	0	7	50	43	0	0	36	36	27
10	0	1	22	44	32	0	2	16	54	29	0	7	7	50	36	0	9	45	36	9
11	0	2	24	47	24	0	2	24	55	20	0	7	14	50	29	0	18	18	27	36
12	0	0	8	40	53	0	2	7	36	55	0	0	0	29	71	0	0	0	63	36

SD = Strongly Disagree D = Disagree S = Sometimes A = Agree SA = Strongly Agree

The principal estimates of the reduced class size effect on teaching present the same picture of teaching as revealed by the classroom studies and the Teacher Questionnaire. Individualization, diagnosis of student strengths and weaknesses, treatment of learning problems, assessment of progress, immediate feedback, and an environment or human relationships conducive to learning were frequently mentioned. One principal comments in the following way:

> *Students are provided with a great deal more individualized attention and assistance as needed, on a regular basis. This can be counted on as the norm.*

Another principal comments:

> *The major changes in teaching that have occurred at …as a result of smaller class sizes are students becoming more engaged in learning, and teachers becoming more engaged in teaching.*

A principal from another school states:

> *I believe that learning has increased as well as teachers' teaching. The smaller classes have enabled the teachers to provide more meaningful activities and spend more time teaching rather than disciplining and setting the atmosphere for teaching.*

A number of principals commented on the effect of SAGE on students with special educational needs. Excerpts from two principals' comments illustrate a perceived trend of classroom teachers being able to attend to the needs of these students:

> *More interventions are being done by the classroom teacher rather than referring them out to the "specialists" in the building.*
>
> *…because our building does not offer any special education classes, parents opt to keep their children here rather than transporting them to the main elementary school when they are tested and found needing services. The smaller number allows these students to achieve at grade level.*

In the Teacher Questionnaire, teachers also report their perceptions about student participation. Ninety percent of the teachers agree or strongly agree that students *participate more in class* and *are more apt to ask for help.* Eighty-five percent of the teachers see their students as *more attentive* and *more enthusiastic about tasks*. And 75% of the teachers see their students *displaying more self-direction*.

Rigorous Curriculum

For the purposes of the SAGE evaluation, rigorous curriculum has been defined as curriculum that is consistent with national standards in reading, language arts and mathematics as proposed by professional associations. Table 46 reports the extent to which the curriculum in the areas of reading, language arts and mathematics in SAGE schools is consistent with these standards. These data, derived from teacher perceptions on the Teacher Questionnaire, show overall agreement with the standards in both curriculum areas. In reading and language arts, the areas of greatest agreement are a) students are encouraged to choose books of personal interest, b) the names of parts of books are taught, c) students are taught to apply a variety of decoding strategies, and d) students are introduced to text that deal with topics relevant to the real world. The areas of least agreement are a) students are taught to critique non-print media, b) students are taught to critique print texts, c) students are taught to categorize texts by fiction or non-fiction, and d) and students are taught to categorize texts by author.

In mathematics, the areas of greatest agreement are a) students have the opportunity to connect mathematics to everyday situations, b) students learn the enumeration system through concrete experiences, c) students have the opportunity to connect mathematics with other subject areas, and d) mathematical language and symbols are introduced in the context of exploration and are related to students' everyday language. The areas of least agreement are a) instruction that includes concrete experiences with metric units, b) use of calculators in appropriate situations, c) development of own mathematics problems, and d) writing in math class to reflect and demonstrate understanding.

Teacher perceptions concerning rigorous curriculum are very similar to prior years, with greater agreement in reading and language arts than in mathematics. The areas of greatest and

least agreement within both curricular areas are nearly identical to previous year's findings.

Table 46. Rigorous Curriculum, Grades K-3 (N, Mean, and Standard Deviation)

	N	Mean*	SD
Reading/Language Arts			
Students introduced to texts: represent range of genres	403	4.14	.68
Students introduced to texts: represent range of historical	404	3.54	.79
Students introduced to texts: deal with topics relevant to real world	406	4.11	.62
Students introduced to texts: variety of ethnic, culture contexts	403	4.03	.67
Students taught to apply variety of decoding strategies	405	4.57	.58
Students introduced to variety of interpretative strategies	402	3.80	.82
Students taught names for parts of books	408	4.70	.53
Students introduced to literature terminology	408	4.47	.75
Students taught to categorize texts: fiction or non-fiction	406	4.11	.94
Students taught to categorize texts: topic or theme	404	3.89	.80
Students taught to categorize texts: author	405	3.85	.87
Students taught to make associations among texts	407	3.8	.71
Student taught aware of how language can be purpose adjusted	402	3.67	.77
Students taught aware of how language can be audience adjusted	405	3.42	.82
Students encouraged to choose books interested in reading	402	4.71	.50
Students apply lang/conventions: critique/discuss print texts	394	3.61	.99
Students apply lang/conventions: critique/discuss non-print media	395	3.13	.90
Students apply lang/conventions: writing to develop interests	398	4.09	.74
Students apply lang/conventions: speaking to develop interests	396	3.79	.84
Mathematics			
Students write own mathematics problem about real or imaginary	406	3.28	.86
Students encouraged to develop own strategy for solving problems	407	4.08	.76
Opportunity to investigate open problems have more than one sol.	407	3.77	.81
Write in math class to reflect and demonstrate understanding	403	3.48	1.00
Mathematics language and symbols introduced in context of explorations	406	4.15	.69
Opportunities to make connections between mathematics and other	407	4.04	.64
Opportunities to make connections between math & everyday	407	4.22	.62
Estimation when working with quantities, measurement, computation	407	3.77	.75
Opportunity to explore and use estimation strategies in real situations	407	3.59	.76
Learn enumeration through concrete experiences	406	4.18	.67
Discuss, model, draw, write about their understanding	407	3.93	.82
Instruction of facts emphasize development of thinking strategies	404	4.10	.71
Develop own computation strategies and algorithms	399	3.62	1.03
Calculators used in appropriate situations	401	2.71	1.17
Instruction includes concrete experiences with metric units	403	3.36	2.25
Concepts of perimeter, area, volume are developed	400	3.38	.91
Opportunity to explore geometric shapes through concrete exp.	406	3.89	.72
Opportunity to work with 3-dimensional figures	407	3.63	.82
Formulate & solve problems involving collecting & analyzing data	406	3.63	.79
Make predictions, inferences, decisions from data	407	3.79	.73
Concept of chance explored by collection of data and other events	405	3.23	.83
Concrete and real experience to develop fraction concepts	405	3.68	.87
Recognize, describe, extend patterns	407	4.17	.71
Create patterns using materials and discuss patterns	407	4.03	.80

*Mean score using five point Likert Scale

The Principal Questionnaire results support the finding that the reading and language arts curriculum and the mathematics curriculum generally are consistent with national standards. All of the SAGE principals see their reading/language arts curriculum as being mostly or completely compliant in these areas, as seen in Table 47. About 93% of the principals regard their mathematics curriculum as mostly or completely compliant with national standards.

Table 47. Principal's Perceptions of Rigorous Academic Curriculum (Percentages)

	Not Implemented	Somewhat Implemented	Mostly Implemented	Completely Implemented
Area				
Reading/ Language Arts	0	0	62.1	37.9
Mathematics	0	6.9	55.2	37.9

N=29

Professional Development

Results concerning general and personal professional development as perceived by SAGE principals and teachers are contained in Tables 48, 49 and 50. Principals' views of the professional development program in their schools are reported in Table 48. The results show that new teacher transitions, collaborative planning, professional development, and staff evaluation programs generally are being implemented in SAGE schools.

Table 48. Principal's Perceptions of Staff Professional Development Programs (Percentages)

	Not Implemented	Somewhat Implemented	Mostly Implemented	Completely Implemented
New teacher transition program	0	10.3	58.6	31.0
Collaborative planning	0	6.9	44.8	48.3
Professional development plans	0	6.9	58.6	34.5
Staff evaluation program	0	6.9	37.9	55.2

N=29

Table 49, which reports the context, process, and content of professional development in SAGE schools, shows that professional development is a prominent feature of SAGE schools. In terms of context, most teachers agree that in their school staff development is an ongoing and regular component, is widely supported, adequately funded, and brings about changes in classroom practices. In terms of process, most teachers agree that in their school the learning climate of staff development is collaborative; the teacher is seen as a learner; and the school's improvement plan addresses decision making, communication, and team functioning. In terms of content of professional development at their schools, teachers report high agreement in the area of child learning and development; knowledge, attitude, and skills needed for quality education; knowledge of effective approaches to teaching; use of strategies that demonstrate high expectations; and a focus on student achievement as a goal and performance assessment.

Areas of professional development in which there is some disagreement by teachers are the use of study groups to learn about change and innovations; out-of-school collaborative learning; learning about innovations prior to deciding about their use; assessing teachers based on student learning; and development activities that include theory as well as practice. Most teachers indicate a lack of staff development activities that specifically target teaching strategies for reduced size classes.

As indicated by findings from the analysis of teaching behaviors in highly effective and less effective SAGE classrooms, differences in emphasis of teaching strategies and classroom management exist among teachers in the classrooms identified for this study. The differences are displayed in Figure 2 (A Model of Less Effective Reduced Class Size Teaching) and Figure 3 (A Model of More Effective Reduced Class Size Teaching) in this report. These early findings will be examined further in next year's study of teaching behaviors and could be instrumental in the guidance and expansion of professional development activities for teachers in reduced size classes.

Table 49. Teachers' Perceptions of Professional Development Grades K-3 (Percentages)

Item	Strongly Disagree	Disagree	Sometimes	Agree	Strongly Agree
1. Ongoing & Regular	0.7	2.9	14.4	45.9	34.1
2. Changes in Practice	1.0	3.7	32.2	47.1	14.6
3. Adequate Funding	4.4	15.9	26.8	36.8	14.6
4. Widespread Support	2.0	10.5	28.0	42.7	15.4
5. Joint Learning	11.0	35.1	32.2	16.1	3.9
6. Study Groups	8.8	24.4	32.9	26.3	6.1
7. Improvement Plan	1.5	7.6	30.7	49.3	8.3
8. "Teacher as Learner"	0.7	6.3	27.3	51.0	12.7
9. Staff Development	1.0	3.9	25.4	53.2	15.1
10. Precede Decisions	2.4	11.2	48.5	31.7	4.6
11. Program Evaluation	1.5	11.2	43.9	37.1	3.9
12. Staff Development Activities	1.7	14.6	44.1	32.4	5.6
13. Teachers Knowledgeable	0.7	05	13.7	55.1	28.5
14. Ensure Quality	0	0.7	9.0	48.8	40.5
15. Effective Approaches	0	1.0	15.9	54.1	28.0
16. Strategies	0	0.2	15.6	50.7	32.4
17. Focus on Goals & Curriculum	0.2	2.2	23.2	51.7	21.5
18. Performance Assessments	1.0	2.2	24.4	53.7	17.8
19. Staff Development for reduced class sizes	4.1	13.2	42.2	30.2	8.8

N=410

Teacher views of their own professional development, as reported in Table 50, show that teachers are divided nearly evenly according to those who have and those who do not have a personal, written development plan. For those who have a personal development plan, in almost all cases, it is developed by the teachers themselves or in consultation with a school administrator. The results also show that most teachers collaborate in planning activities, delivering lessons, evaluating students, and in school-wide instructional initiatives. Further, teachers attend conferences and take improvement courses. Although participation increased over last year, few SAGE teachers, when compared to other activities, attend a workshop or seminar on teaching small classes. This is probably the case because courses with this focus may not be available at the present time.

Table 50. Teachers' Perceptions of Their Personal Professional Development (N=410)

Question #20 ***Over the past year, I have…***	Percentage of responses
Engaged in a mentoring relationship with another teacher.	40.7
Participated in joint planning activities with other SAGE teachers.	91.7
Collaborated with other teachers in delivering lessons.	84.6
Collaborated with other teachers in evaluating student progress.	87.8
Participated in a study group or on-line network.	31.5
Collaborated in school-wide instructional initiatives or themes.	72.2
Collaborated with other schools or institutions.	41.5
Conducted research connected to my teaching.	36.3
Attended a professional conference or skill-building workshop.	79.8
Attended a workshop, seminar or retreat focused on diversity or human relations training.	30.5
Attended a workshop, seminar or retreat focused on teaching smaller classes.	20.5
Taken a course for graduate of CEU credit.	55.6
Question 21 ***Do you have a personal formal, written professional development plan?***	
Yes	48.8
No	49.5
Question 22 ***Which of the following statements most accurately reflects the content of your professional development plan?***	
It was determined primarily by me	33.9
It was determined in consultation with school administrators.	12.9
It was determined in consultations with district administrators.	1.2
It was determined primarily by school and /or district administrators	4.9

Family Involvement and Lighted Schoolhouse

The extent to which SAGE school parents are involved in education of their children is reported in Table 51. The results of the Teacher Questionnaire show that teacher-parent contacts occur mostly through teacher notes, teacher and parent conversations, and telephone calls. The use of weekly progress reports requiring a parent signature is increasing, however. The use of weekly progress reports in SAGE schools has more than doubled compared to its reported use in 1996-97 and in 1997-98.

Table 51. Teacher Questionnaire Results for Family Involvement (Percentages)

Item	1996-97 (N=212)	1997-98 (N=315)	1998-1999 (N=417)	1999-00 (N=410)
Class Newsletter	71	62	62	64
Weekly progress report-requiring parent signature	24	28	54	62
Weekly progress report-not requiring parent sig.	11	12	50	48
Notes sent home	98	93	95	95
Conversations with parents	95	94	95	97
Parental visits to school	74	76	71	74
Telephone calls	92	89	91	94
Home visits	10	14	12	13

School-wide opportunities for family involvement reported by principals are shown in Table 52. The 1998-99 data on lighted schoolhouse activities were collected directly from the school principals rather than through the Department of Public Instruction, as had been the case in previous years. The data in Table 52 are based on 29 completed questionnaires.

Table 52. SAGE Schools' Lighted Schoolhouse Participation (N=29)

Activity	Number of Schools Reporting the Activity	Range of Participants in Each Activity	Total Number of Annual Participants
Child Care	11	20-200	761
Health Clinic	9	10-580	1420
Breakfast	25	15-460	3744
Tutoring	25	4-500	2208
Homework Help	17	8-150	885
Extended Library	12	15-1200	2961
Adult Recreation	19	15-1200	2784
Girl and Boy Scouts	28	6-360	1850
Music Lessons	12	2-180	529
Summer Reading	23	15-250	2097
Head Start	5	3-100	224
Social Services	6	12-720	1227
Family Resource Center	12	5-292	997
Technology Education	7	20-580	1015
GED Preparation	2	20-30	50
PTA/PTO	22	5-800	2315
Family Literacy	8	17-200	458
Parent Advisory	15	4-25	184

Principals also reported a number of additional activities well attended by SAGE families, such as meal activities which varied from Sunday brunches, chili dinners, spaghetti dinners, holiday dinners, and invitations to SAGE parents once a month for the routine student breakfasts. A variety of special activities such as family fun days, reading nights, career exploration days, 4H clubs, open house activities, science fairs, and theatre productions were also reported.

MAJOR FINDINGS AND DISCUSSION 1999-2000

The Student Achievement Guarantee in Education (SAGE) program is a statewide effort to increase the academic achievement of children living in poverty by reducing the student-teacher ratio in kindergarten through third grade to 15:1. Schools participating in the SAGE program are also required to implement a rigorous academic curriculum, to provide before- and after-school activities for both students and community members, and to implement professional development and accountability plans. The SAGE evaluation is being conducted under contract with the Department of Public Instruction by the School of Education at the University of Wisconsin–Milwaukee.

During the 1996–97 school year SAGE was implemented in 30 schools located in 21 school districts. It encompassed 84 kindergarten classrooms, 96 first grade classrooms, and 5 mixed grade classrooms enrolling 1,715 kindergarten and 1,899 first grade students. In 1997-98, the SAGE evaluation added 113 second grade classrooms in the original 30 SAGE schools. In 1998-99, the SAGE evaluation was made up of 131 kindergarten, 143 first grade, 143 second grade and 139 third grade classrooms enrolling 2,303 kindergarten, 2,508 first grade, 2,493 second grade and 2,572 third grade students. In 1999-00, kindergarten and first grade students were not evaluated. The 1999-00 SAGE evaluation was made up of 89 second grade and 83 third grade classrooms enrolling 2,624 and 2,656 students respectively.

To measure academic achievement, second- and third-grade students in SAGE schools and in a group of Comparison schools were administered the Comprehensive Test of Basic Skills (CTBS) Complete Battery, Terra Nova edition, Levels 12 (2nd grade) and 13 (3rd grade) in the spring of 2000. Following is a summary of the major findings.

The Achievement Effect of Class Size Reduction

Second grade, 1999-00

- When adjusted for pre-existing differences, second grade SAGE students showed a significant achievement advantage over their Comparison group counterparts in all areas. The exception was in the area of reading when first grade post-test was used to adjust for achievement differences (Table 17).
- African American SAGE second graders scored significantly higher than African American Comparison school students in mathematics, language arts, and total scores at the end of second grade. When examining gains made in 1998-00 from the first grade pre-test to the second grade, African American SAGE students made significantly larger gains than their Comparison school counterparts on the total scale score as well as every sub-test except for language arts. When using the first grade post-test as the baseline, African American SAGE students again made larger gains in 1999-00 on every test except for reading, but the gains were not statistically significant (Table 19).
- African American students, as a group, scored significantly lower than white students on total scale scores and on all sub-tests, regardless of whether they were in SAGE or Comparison schools, although the gap between African Americans and whites is larger in Comparison schools (Table 20).

Third Grade, 1999-00

- When adjusted for pre-existing differences in academic achievement, attendance, socioeconomic status and race, SAGE students showed significant improvement over

their Comparison school counterparts from the beginning of first grade to the end of third grade across all academic areas (Table 27). From the beginning of second grade (first grade post-test) significant additional differences in gain were seen in mathematics (Table 28). From the beginning of third grade no significant additional differences in gain advantage were found (Table 29).

- African American students continued to score significantly lower than white students on total scale score and on all sub-tests, regardless of whether they were SAGE or Comparison school students. Gains made by African American versus white students were significantly better in SAGE schools from the beginning of first grade to the end of third grade. The opposite pattern was observed in Comparison schools (Table 32).

Discussion of the Achievement Effect

Analyses of SAGE achievement test results suggest that, overall, first grade achievement gains are significantly higher for SAGE students than for Comparison school students. SAGE students appear to retain this advantage in second and third grades.

SAGE African-American students narrow the achievement gap with white students in first grade and keep pace with white students in second and third grade.

One factor not yet examined is the extent to which students with no SAGE experience entering SAGE classrooms for the first time in second and in third grade influence the achievement results of SAGE second and third grade classrooms. The potential impact of these new students merits investigation.

The Analysis of SAGE Classrooms and Schools

Data collected from interviews and observations of select first-grade SAGE classrooms in 1999-00 along with findings from the teacher questionnaire administered to all SAGE teachers and principal questionnaire completed by all SAGE principals clarify and extend SAGE results from 1996-97, 1997-98, and 1998-99 regarding classroom events.

- The finding from previous years that teaching in reduced size classrooms is characterized by more individualization, time spent on teaching rather than disciplining, class discussion, hands on activities, content coverage, and teacher enthusiasm is reinforced.
- The degree of individualization common to all SAGE classrooms is intensified in higher achieving first-grade classrooms.
- The individualization in higher-achieving classrooms is teacher-centered and characterized by teacher behavior that encompasses large amounts of time spent on the monitoring of learning, the eliciting of understandings, requiring students to display knowledge and skills, providing feedback and critique, and re-teaching when necessary.
- First-grade teachers in higher-achieving classrooms emphasize basic skills and processes through modeling, drill, and practice.
- First-grade teachers in higher achieving classrooms prefer highly structured, goal-directed classrooms with established routines where learning proceeds at a quick pace.

- Classroom management of teachers in higher achieving classrooms is firm and decisive, but also positive and nurturing.
- First-grade teachers of lower achieving classrooms tend to believe that the primary advantages of a reduced size class are the opportunity to develop critical thinking, to permit students to choose their activities, and to implement more activities and problem solving lessons.

Discussion

Taken together, analyses of teaching in SAGE higher-performing classrooms suggest that SAGE schools should focus their staff development efforts on helping teachers learn to increase the focus of their instruction on academic learning, employ teacher directed basics oriented individualization with special emphasis on student articulation of understandings, teacher critique and re-teaching.

References

Allen, R. (1997). Effects of the revenue caps on Wisconsin's school districts. 1996-97 school year fourth year study. Madison, WI: Wisconsin Education Association Council.

Bingham, S. C. (1993). *White-minority achievement gap reduction and small class size: A research and literature review*. Nashville, TN: Center of Excellence for Research and Policy on Basic Skills.

Boyd-Zaharias, J. and Pate-Bain, H. (2000). Early and new findings from Tennessee's project STAR. *The CEIC Review, 9*(2), 4.

Bryk, A. & Raudenbush, S. (1992). *Hierarchical linear models*. Newbury Park, CA: SAGE Publications, Inc.

CTB/McGraw-Hill (1991). *Comprehensive test of basic skills* (4th ed.). CA: Macmillan/McGraw-Hill.

Cohen, J. (1977). *Statistical Power Analysis for the Behavioral Sciences* (rev. ed.). New York: Academic Press.

Finn, J. D., Gerber, S. B, & Farber, S. F. (2000). Teacher aides: An alternative to small classes? *The CEIC Review, 9*(2), 5.

Glass, G., & Smith, M. (1978). Meta-analysis of research on relationship of class-size and achievement. San Francisco, CA: Far West Laboratory for Educational Research and Development.

Grissmer, D. (1999). Class size effects: Assessing the evidence, its policy implications, and future research agenda. *Educational Evaluation and Policy Analysis*, *21*(2), 231-248.

Hanushek, E. A. (1999). Some findings form an independent investigation of the Tennessee STAR experiment and form other investigations of class size effects. *Educational Evaluation and Policy Analysis, 21*(2), 143-163.

Hruz, T. (1998). Beyond smoke and mirrors: A critical look at smaller class sizes. *Wisconsin Interest, 7*(2), 29-37.

Hruz, T. (1999). The Costs and Benefits of Smaller Classes in Wisconsin. Wisconsin Policy Research Institute.

Krueger, A. B. (2000). An Economist's view of class size research. *The CEIC Review, 9*(2), 19-20.

Mosteller, F. (1995). The Tennessee study of class size in the early school grades. *The Future of Children, 5,* 113-127.

Pardini, P. (1998). Class wars. *Middle Ground, 1*(4), 22-24.

Pate-Bain, H., Achilles, C. M., Boyd-Zaharias, J., & McKenna, B. (1992). Class size does make a difference. *Phi Delta Kappan, 74*(3), 253-256.

U. S. Department of Education and the Laboratory for Student Success, (1999). How small classes help teachers do their best. Commissioned papers for the National Invitational Conference "How Small Classes Help Teachers Do Their Best" held on December 6-7, 1999 in Washington, DC.

Viadero, D. (1998). Small classes: Popular, but still unproven. *Education Week.* Available: http://www.edweek.org/ew/vol-17/23class.h17.

Wenglinsky, H. (1997). *When money matters.* Princeton, NJ: Educational Testing Service.

Appendix A

Reading/Mathematics Observations

Directions

1. Observe in *reading* (not language arts or separate phonics) and *mathematics*.

2. Observe for the complete class session.

3. Obtain copies of all distributed photocopied papers, directions, materials, etc.

4. Obtain titles, etc., of all published materials used.

5. Informally interview the teacher if clarification about the observations is needed.

Guide

What's going on? Describe the events of the lesson in the order that they occur. Focus on the following:

Learning activities
Teacher and student behavior and talk (verbatim)
Teacher role, methods, techniques, etc.
Student role, who participates, type of participation
Objectives (stated or inferred)
Specific content/skills being taught
Materials and resources
Class organization: grouping, setting, etc.
Evaluation

2. What's going on in terms of previous SAGE findings? Describe the extent to which each of the following does or does not occur:

Individualization (one-to-one, small group, active participation in total class0

Disciplinary or class structuring behavior (reprimanding, praising, etc.)

Hands-on active student involvement activities (manipulatives, problem solving, creative tasks)

Content beyond the grade level or deeper, extended content (critical thinking, second-grade curriculum)

High academic learning time (students engaged in content or skill learning, little time spent in giving directions, collecting papers, etc.)

Enthusiasm for teaching (teacher excitement, energy, effort, etc.)

Appendix B

1999-2000 SAGE Classroom Studies
Introductory Interview

As the SAGE Evaluation continues into its fourth year, we want to come to understand better how class size affects teaching at the first-grade level. In 1999-2000 we are focusing on SAGE teachers who have at least two years of experience in the program.

1. Describe how students are assigned to classrooms in your school.

   Probes:
   - EEN
   - Parent requests
   - Student ability
   - Behavior
   - Random

2. Describe what you believe are two or three more dominant features of your teaching. That is, what would a colleague, parent, or someone who knew your work say characterizes your teaching? Describe your methods.

3. Describe what, if anything, is new to your teaching this year. What are you doing this year in your first grade that you did not do last year?
   Probes:
   - Methods,
   - Curriculum/content
   - Materials
   - Organization

4. Talk about test administration.
   Probes:
   - How do you prepare your students for testing?
   - Academic preparation?
   - Emotional and physical wellbeing on testing day?
   - What time of day do students usually take the test?
   - What is your role in the testing situation (level of support)?
   - What are your feelings about testing in general?
   - Do your students do well on SAGE testing?

Appendix C

1999-2000 SAGE Classroom Studies
Reading Interview

Teacher: _____
Date: _____

1. Describe your philosophy of reading instruction in your first grade.

   Probes:

   - Goal/Objective (decoding skills, comprehension, interest in reading, etc.)
   - Curricular program (whole language, phonics, etc.)
   - Teaching methods (direct instruction – teacher centered, workshop – student centered, etc.)
   - Materials (basal readers, program texts, trade books, etc.)

2. Describe a typical reading lesson or session in your (reduced-size) first grade. Tell what you do or what happens first, second, and so on.

   Probes:

   - Goal/Objective
   - Curricular program
   - Teaching methods
   - Materials
   - Evaluation (oral reading, monitoring, tests, formal MPS or school evaluations etc.)
   - Helpers (aides, parent helpers)
   - **Time** (lesson duration, minutes per week, actual time spent reading)
   - Individualization
     Grouping (number, homo vs. hetero)
     One-to-one

3. Describe what, if anything, would improve reading instruction in your classroom.

Appendix D

1999-2000 SAGE Classroom Studies
Mathematics Interview

Teacher: _____
Date: _____

1. Describe your philosophy of mathematics instruction in your first grade.
   Probes:

   - Goal/Objective (conceptual understanding, computation skills, problem solving, etc.)
   - Curricular Program
   - Teaching methods (direct instruction – teacher centered, manipulatives and discovery – student centered)
   - Materials (textbooks, workbooks, manipulatives, etc.)

2. Describe a typical mathematics lesson or session in your (reduced-size) first grade. Tell what you do or what happens first, second, third, and so on.

   Probes:
   - Goal/Objective
   - Curricular Program
   - Teaching methods
   - Materials
   - Evaluation (timed fact tests, problem-solving tests, processes used)
   - Helpers (aides, parent helpers)
   - **Time** (lesson duration, minutes per week)
   - Individualization
     Grouping (number, homo vs. hetero)
     One-to-one help

3. Describe what, if anything, would improve mathematics instruction in your first grade.

Appendix E

1999-2000 SAGE Classroom Studies
Self Report

Teacher Background Information:

Teacher: _____ Class enrollment: 99-00 ____
98-99 ____
97-98 ____

Type of SAGE Classroom: _____ School: _____

1. Teacher Background: Certification _____Majors/Minors_____

   Reading License: _____Degrees: _____

   Inservice/Workshops:_____

   _____

   Other:_____

   _____

2. Teaching Experience: _____

   How many years in first grade? ____ How many years at this school? ____

3. Other adults who help with math or reading instruction in your classroom.

Name of adult	Amount of time they spend with students on a daily basis in Reading	in Math
_____	_____	_____
_____	_____	_____
_____	_____	_____

Survey - 1999-2000 SAGE Classroom Studies

SAGE Evaluation is now focusing on teachers who have at least two years experience in the program. Below are questions designed to determine how SAGE teachers allot their time and also where they direct their efforts.

1. A good teacher adjusts instruction according to the particulars of the class. Thinking about such student qualities as interests, abilities, needs and personalities, would you say your teaching leans more toward the particulars of the students as individuals or more toward the particulars of the class in general?

   students as individuals……class in general
   1 2 3 4 5

2. Teachers vary in enthusiasm from year to year. How would you characterize your enthusiasm for teaching?

   a. Last year?
   not enthusiastic……very enthusiastic
   1 2 3 4 5

   b. This year?
   not enthusiastic……very enthusiastic
   1 2 3 4 5

   c. Compared to other teachers in the building?
   not enthusiastic……very enthusiastic
   1 2 3 4 5

3. Teachers make decisions about their approach to curriculum content, specifically about its depth and breadth.

   a. Thinking about math content, how would you estimate your students' depth of understanding at the end of last year?
   not deep……very deep
   1 2 3 4 5

   b. How much math content do you typically cover in a year?
   cover the grade level content ……...work into the next grade level
   1 2 3 4 5

   c. Thinking about reading content, what would you estimate to be your students' depth of understanding at the end of last year?

not deep……very deep
1 2 3 4 5

d. How much reading content do you typically cover in a year?
cover the grade level content ……...work into the next grade level
1 2 3 4 5

4. Budgeting time is another dilemma for teachers. Typically, how much time do you spend on each of these tasks?

a. Diagnosing the needs of individual students?

A very little amount of time……...A very large amount of time
1 2 3 4 5

b. Disciplining the class?
1 2 3 4 5

c. Providing help for individual students?
1 2 3 4 5

d. Working with small groups of students?
1 2 3 4 5

e. Assessing the progress of individual students?
1 2 3 4 5

f. Assessing the progress of the class in general?
1 2 3 4 5

g. Working with students one-on-one?
1 2 3 4 5

h. Disciplining individual students?
1 2 3 4 5

i. Engaging students in hands-on activities?
1 2 3 4 5

j. Working with students on special projects?
1 2 3 4 5

k. Developing creative projects for your class?

1 2 3 4 5

l. Using direct instruction, meaning that you explain and give information to the class, model, practice, provide feedback, etc.

1 2 3 4 5

Indicate the two areas from a.) through l.)on which you wish you could spend more time.