1997-98 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

Lee Breese
Karen Ehrle
Anke Halbach
Amanda Palmer
Alan Silverman

Project Administrator

William Harvey

DECEMBER 1998

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/SOE/centers&projects/sage/

TABLE OF CONTENTS

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 SAGE statute [s. 118.43] requires 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 14-16 non-SAGE comparison schools located in SAGE districts. The results of the 1996-97 and 1997-98 evaluations are 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. However, it is worth noting two significant design differences. First, the Tennessee STAR Project used a true experimental design. The SAGE project uses a quasi-experimental design.

This has two implications. The SAGE project evaluation uses naturally occurring classrooms (the most realistic conditions) while STAR employed random assignment of students to classroom types which were held constant for the duration of the study. Secondly, 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:

<u>SAGE vs. Comparison School – Achievement Outcome Questions</u>

1. What are the achievement levels of SAGE classrooms compared to achievement levels of classrooms in selected comparison schools?
2. Does attendance in a SAGE classroom have a differential impact on the achievement of minority students and white students?
3. 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?
4. Does the impact on achievement of participation in a SAGE classroom change from year to year as students progress from first through third grade?

<u>SAGE Schools – Classroom and School Questions</u>

1. What are the instructional characteristics of SAGE classrooms?
2. How are SAGE classrooms organized?
3. Does the type and extent of student participation in SAGE classrooms correlate with achievement outcomes?
4. Does the style of teaching in SAGE classrooms correlate with achievement outcomes?
5. 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?
6. 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?
7. What is the structure and focus of professional development activities in SAGE schools?
8. 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 been an increasingly popular policy in recent years. 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 an aide, or 25 with no aide. 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 benefited, disadvantaged minority students seemed to benefit more from smaller class sizes than their peers (Pate-Bain, Achilles, Boyd-Zaharias, & McKenna, 1992).

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.

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. 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).

In Wisconsin, most class sizes exceed the class size standards set by the STAR Project. A study completed by Allen (1997) of K-6 class sizes in Wisconsin's public schools reported that 92% of Wisconsin's kindergarten classes exceed the lower class size standard established by

Project STAR of 15 or fewer students per teacher. Twenty-seven percent of Wisconsin kindergarten teachers reported class sizes that exceeded 25 students per teacher.

Summary of 1996-97 Findings

Achievement Outcome Findings 1996-97

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 the CTBS 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, income, 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. On averaged classroom scores, 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 data on the average performance of students in SAGE classrooms suggest that the lower student-teacher ratio in SAGE classrooms mitigates 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.

School and Classroom Findings 1996-97

To more fully understand the SAGE program, it is important to understand how SAGE schools structured classrooms and implemented a reduced student-teacher ratio, rigorous curriculum, staff development, and lighted schoolhouse. Together, that information provides a within SAGE school and classroom data description of life in SAGE classrooms and schools.

Classroom Level Findings 1996-97

Data from 1996-97 suggested that after one year of the SAGE program classroom discipline problems are greatly reduced, and when classroom management was needed, it was overwhelmingly positive. The direct benefit of having to spend less time managing the class was increased instructional time, i.e., actual time spent on teaching. Further, the increased instructional time available to teachers was used to attend to the learning needs of individual students.

The type of instruction that students encountered in SAGE classrooms was predominantly teacher centered. Listening, practicing, receiving help, and answering accounted for between 50 to 75 percent of the teaching-learning that occurred. Although teachers indicate that their use of more student-centered activities such as creating, manipulating, and problem

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

Several anticipated outcomes did not emerge from a composite of the interviews, observations, logs, and questionnaires in 1996-97. While all teachers reported some changes in their teaching, a large swing to student-centered teaching, a change that some might expect as a result of reduced class size, was not observed. Lastly, a clear difference in teaching and learning among the four main types of SAGE classrooms was not apparent.

School Level Findings 1996-97

The Teacher Questionnaire and Principal Interviews, both completed in May 1997, 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 mixed-grade teacher responses indicated that their reading/language arts curricula were more congruent with professional standards than SAGE kindergarten curricula. Teacher responses suggested no important differences in the degree to which their curricula were congruent with professional standards in the area of mathematics. Principal responses to curriculum-related questions suggested that a rigorous curriculum included basic skills, problem solving, and higher-level thinking. Only a handful of principals seemed to believe that the curriculum of their school was rigorous. However, most SAGE principals regarded parts of their curriculum as strong.

A section of the Teacher Questionnaire contained *staff development* questions. Teachers are asked about their individual level of professional development as well as the extent to which their school district provided staff development programs. Despite contractual requirements, roughly 60 percent responded that they had no "personal, formal, written development plan." Twenty-one percent rated their school district's staff development program at the "initialization" phase, about 66 percent rated it at the "implementation" stage, and about 9 percent felt the district was "institutionalizing" the staff development program.

Data regarding implementation of *lighted schoolhouse* activities for 1996-97 were obtained from the Principal Interviews and year-end reports required by DPI. In addition, data regarding lighted schoolhouse activities existing prior to schools' participation in SAGE were obtained from the Baseline Data Questionnaire administered in May 1996 and the school contracts completed for DPI prior to enrollment in the SAGE program. Principal Interview data suggested that because of SAGE schools took responsibility for the conception and operation of the lighted schoolhouse activities. However, the schools tended not to focus heavily on their lighted schoolhouse activities in the first year of SAGE implementation.

EVALUATION DESIGN .

Descriptions and Definitions

Schools

During 1997–98, the SAGE program was implemented in 30 schools located in 21 school districts throughout the state, as shown in Table 1. In addition, the SAGE program consisted of 14 Comparison schools located in 7 school districts.

Table 1. SAGE Schools 1997-98

SAGE DISTRICTS AND SCHOOLS	
DISTRICT School	**DISTRICT** School
ADAMS-FRIENDSHIP AREA Adams Elementary	**MENOMONEE 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 of Lifelong Learning
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	**SOUTH SHORE** South Shore Elementary (Port Wing)
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 1997-98, the SAGE evaluation involved 7,161 active students in 117 kindergarten, 118 first-grade, and 113 second-grade classrooms. The gender, race, and other characteristics of students in SAGE schools are displayed in Table 2:

Table 2. Characteristics of SAGE Students 1996-97 and 1997-98

Characteristic	Percent of Students 1996-97	Percent of Students 1997-98
Gender		
Female	48.6	49.3
Male	51.4	50.7
Race/Ethnicity		
African American	24.8	24.3
Asian	5.7	5.3
Hispanic	6.6	7.8
Native American	11.7	7.9
White	48.8	46.9
Other	1.6	2.0
Subsidized Lunch Eligibility		
Free	57.7	67.4
Reduced	10.9	11.0
Not Eligible	31.4	21.7
Repeating Grade	3.2	3.2
English as Second Language	8.2	7.1
Referred to M-Team	13.6	13.1
Exceptional Education Need	13.1	13.2
Individualized Education Plan	8.2	7.6

During the course of the 1997-98 school year, records were compiled on 8,843 students. Many students withdrew from SAGE and comparison schools during the year, while others enrolled. Those students who remained in their schools for the entire year are labeled "persisters". As Table 3 shows, enrollment in comparison schools was slightly more stable than in SAGE schools. Moreover, in both SAGE and comparison schools, the number of students withdrawing exceeded the number of students enrolling for the year. Thus, the number of persisters plus newly enrolled students recorded during spring data collection totals 7161, distributed across schools and grades as shown in Table 4.

Table 3. Enrollment Changes in SAGE and Comparison Schools by School Year (Number of Students and Percentage of Students)

	SAGE				COMPARISON				ALL SCHOOLS			
	1996-97		1997-98		1996-97		1997-98		1996-97		1997-98	
	#	%	#	%	#	%	#	%	#	%	#	%
Ongoing	2943	81.4	2455	42.3	1706	85.3	1402	44.3	4649	82.8	3857	43
Withdrew	397	11	1093	18.8	178	8.9	589	18.6	575	10.2	1.682	18.7
Enrolled	274	7.6	2262	39	115	5.8	1175	37	389	6.9	3437	38

Table 4. Number of Students in SAGE and Comparison Schools by Grade and School Year

	SAGE		COMPARISON		TOTAL	
	1996-97	1997-98	1996-97	1997-98	1996-97	1997-98
Kindergarten	1494	1524	820	676	2314	2200
first Grade	1723	1567	1001	985	2724	2552
second Grade	NA	1541	NA	868	NA	2409
Totals	3217	4632	1821	2529	5038	7161

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.

Two other types of classroom organization were also utilized in the SAGE program, but to a limited extent. They are the *Split Day* classroom consisting of 15 students and two teachers, one who teaches in the morning and one who teaches in the afternoon, and the *3-Teacher Team* classroom where there are 45 students taught collaboratively by three teachers.

The types of classrooms are displayed in Table 5. SAGE classes range in number of students from 7 to 38. A few SAGE classrooms exceed the 15:1 student-teacher ratio, but only by a few students.

Table 5. Number of SAGE Classrooms by Type, Grade, and School Year

	Regular		**2-Teacher Team**		**Floating Teacher**		**Shared Space**		**Split Day**		**3-Teacher Team**	
	96-97	**97-98**	**96-97**	**97-98**	**96-97**	**97-98**	**96-97**	**97-98**	**96-97**	**97-98**	**96-97**	**97-98**
Kindergarten	50	89	24	22	3	2	2	4	0	0	1	0
Grade 1	61	84	18	23	7	2	8	8	2	0	0	1
Grade 2	NA	82	NA	21	NA	3	NA	6	NA	0	NA	1

Data Collection Instruments

To provide information about the processes and product of the SAGE program for 1996-97 and 1997–98, 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 and 1997-98 follows. The data collection instruments and the plan for their use throughout the evaluation are displayed in Tables 6 and 7.

1. *Comprehensive Test of Basic Skills (CTBS)*. 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. 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

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

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.

Table 6. Cohort CTBS Testing by Grade Level 1996-01

1996-97	**1997-98**	**1998-99**	**1999-00**	**2000-01**
K Cohort 1	K Cohort 2	K Cohort 3	K	K
1 (fall & spring)	1(fall & spring)	1(fall & spring)	1	1
	2(spring)	2(spring)	2(spring)	2
		3(spring)	3(spring)	3(spring)

2. *Student Profiles.* This instrument completed in October and May, provided demographic and other data on each SAGE school and comparison school student.
3. *Classroom Organization Profile*. Completed in October, this instrument was used to record how SAGE schools attained a 15:1 student-teacher ratio.
4. *Principal Interviews*. 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.

5. *Teacher Questionnaire*. 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*. 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*. In both October and May, teachers used this instrument to assess each student's level of participation in classroom activities.
8. *Classroom Observations.* A group of first-grade and second-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*. Although in-depth teacher interviews were not part of the original SAGE evaluation design, they were added in 1997 because it became apparent that teachers had important stories to tell about their SAGE classroom experiences. The interviews dealt with teachers' perceptions of the effects of SAGE on their teaching and on student learning.

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

	1996–97	**1997–98**	**1998–99**	**1999-2000**	**2000-2001**
Student Participation Questionnaire Fall, Spring	K, 1	K, 1, 2	K, 1, 2, 3	K, 1, 2, 3	K, 1, 2, 3
Teacher Questionnaire Spring	K, 1	K, 1, 2	K, 1, 2, 3	K, 1, 2, 3	K, 1, 2, 3
Teacher Log Fall, Winter, Spring	K, 1	K, 1, 2			
Classroom Observation Fall, Spring	1 (Selected)	1, 2, (Selected)			
Teacher Interview Spring	1 (Selected)	1, 2 (Selected)			
Principal Interview Spring	K, 1	K, 1, 2			
School Case Study Continuous			1, 2, 3 (Selected)	1, 2, 3 (Selected)	1, 2, 3 (Selected)
Principal Questionnaire Spring			K, 1, 2, 3	K, 1, 2, 3	K, 1, 2, 3

ANALYSES OF STUDENT ACHIEVEMENT OUTCOMES 1997-98

Methods Introduction

Statistics Utilized

The 1997-98 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, 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 this 1997-98 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. Finally, hierarchical linear modeling is pertinent to the SAGE evaluation because this technique focuses on the class effects of SAGE; that is, these analyses will 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.

The 1996-97 Report

In its 1996-97 evaluation, the SAGE evaluation team also utilized descriptive statistics and multivariate analyses, including linear regression and hierarchical linear modeling. However, there are two essential differences between the 1997-98 quantitative evaluation and the 1996-97 quantitative evaluation. First, the 1996-97 report included national percentile scores as well as normal curve equivalent scores. National percentile scores are not reported in the 1997-98 summary because the use of national percentile scores in regression analysis is potentially misleading due to the non-equal interval nature of this scale. Instead, normal curve equivalents are included in the descriptive sections of the current report to help clarify the analytical results.

Normal curve equivalents are not reported among the inferential analyses because the results of such analyses would be redundant with those analyses utilizing the scale scores. Second, sections of the 1996-97 report presented analyses based on the exclusion of the top scoring quartile because the post-test given to 1996-97 first graders proved to be too easy, which in essence created a test ceiling effect for top scoring students at this grade level. However, this problem was corrected in the 1997-98 testing with an appropriate post-test level, and therefore the inclusion of these analyses is not necessary (there was no ceiling effect).

General Findings 1996-97

Some general findings from 1996-97 quantitative analysis show 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. As a group, 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. At the individual level of analysis, after controlling for pre-test score, SES, attendance, and race, SAGE first-grade students 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.

Score Metrics 1997-98

A brief discussion of the metrics reported in the 1997-98 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.

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, the focus group, in this example, did not "gain" more than the reference group and thus the score remained constant.

Structure of 1997-98 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: (1) First-Grade Results and (2) Second-Grade Results. The following are delineated 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 modeling.

In addition, the quantitative section includes "within SAGE" analyses for first-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

First-Grade Results 1997-98

Descriptive Statistics

Valid Test Scores. The number of first-grade students for whom the valid test scores are available is substantially less than the total number of students. There are four main explanations for this. First, the evaluation team presented schools with the option of allowing EEN and ESL students to take the test, even though the test may be inappropriate for these students. These scores were invalidated based on a "Nonvalid/Missing Test Report," developed by the evaluation team and completed for all first grade classes. Second, given withdrawals and enrollments during the school year, a number of students had valid pre-test scores, but no post-test scores and vice versa. Third, some students took the reading and language arts components of the CTBS, or the mathematics component, but not both. Consequently, total scores are unavailable for these students. Finally, some of the students did not complete the pre-test, post-test, or both the pre- and post-tests. The number of valid test scores for the 1997-98 school year are presented in Table 8.

Table 8. Number of 1997-98 First-Grade Students with Valid Test Scores

Fall 1997 Pre-Test				**Spring 1998 Post-Test**			
	Total	**SAGE**	**Comparison**		**Total**	**SAGE**	**Comparison**
Reading	2246	1383	863	Reading	2162	1318	844
Language Arts	2245	1383	862	Language Arts	2163	1319	844
Mathematics	2239	1382	857	Mathematics	2175	1334	841
Total	2211	1367	844	Total	2140	1310	829

Pre-Test (Baseline) Results. Table 9 provides descriptive statistics from the pre-test (baseline) results. Both Scale Scores and Normal Curve Equivalents are presented. Given the

longitudinal nature of the SAGE evaluation, scale scores serve as the primary measure of student achievement.

Table 9. Combined SAGE and Comparison Population Descriptive Statistics on CTBS PRE-TEST Results for 1997-98 First-Grade Students

	SCALE SCORES		NORMAL CURVE EQUIVALENT	
	MEAN	**STANDARD DEVIATION**	**MEAN**	**STANDARD DEVIATION**
Reading	533.99	36.31	44.47	19.86
Language Arts	529.84	43.62	43.73	21.34
Mathematics	492.58	41.04	43.28	19.11
Total	519.20	34.59	43.31	19.11

Difference of Means Test. The results from difference of means tests between SAGE and comparison student scale scores from the Fall 1997 CTBS Level 10 Pre-Test are reported in Tables 10-13. Comparison school students scored slightly higher than SAGE school students on the reading sub-test, mathematics sub-test, and total scale, and slightly lower on the language arts sub-test. However, none of these differences is statistically significant at the .05 level. We fail to reject the null hypothesis that there is no difference between SAGE and comparison school students on the pre-test. As a result of SAGE and comparison students essentially being equal in achievement at the beginning of the SAGE program, any differences in the post-test scores benefiting SAGE students may be more assuredly attributed to the student-teacher ratio of 15:1 in the SAGE classroom.

Table 10. Differences of Means Test on Language CTBS FALL PRE-TEST for 1997-98 First-Grade Students

		SCALE SCORES		NORMAL CURVE EQUIVALENTS	
	N	**MEAN**	**STANDARD DEVIATION**	**MEAN**	**STANDARD DEVIATION**
Comparison Schools	862	528.97	43.39	43.25	21.13
SAGE Schools	1383	530.50	43.78	44.08	21.48
*Significant at .05 level					

Table 11. Differences of Means Test on Reading CTBS FALL PRE-TEST for 1997-98 First-Grade Students

		SCALE SCORES		NORMAL CURVE EQUIVALENTS	
	N	**MEAN**	**STANDARD DEVIATION**	**MEAN**	**STANDARD DEVIATION**
Comparison Schools	863	535.06	36.18	45.21	19.10
SAGE Schools	1383	533.35	36.43	44.02	20.33
*Significant at .05 level					

Table 12. Difference of Means Test on Mathematics CTBS FALL PRE-TEST for 1997-98 First-Grade Students

		SCALE SCORES		NORMAL CURVE EQUIVALENTS	
	N	**MEAN**	**STANDARD DEVIATION**	**MEAN**	**STANDARD DEVIATION**
Comparison Schools	857	493.02	38.38	43.36	18.15
SAGE Schools	1382	492.34	42.51	43.25	19.66
*Significant at .05 level					

Table 13. Difference of Means Test on Total CTBS FALL PRE-TEST for 1997-98 First-Grade Students

		SCALE SCORES		NORMAL CURVE EQUIVALENTS	
	N	**MEAN**	**STANDARD DEVIATION**	**MEAN**	**STANDARD DEVIATION**
Comparison Schools	844	519.51	33.35	43.47	18.34
SAGE Schools	1367	519.06	35.34	43.25	19.56
*Significant at .05 level					

As noted above, student populations varied in SAGE and comparison schools due to withdrawals and within-year enrollments. The post-test results are based only on those first-grade students who remained in their schools for the entire 1997-98 school year. CTBS allows for measurement of performance over time and therefore pre-test and post-test scores are comparable from a measurement position. The CTBS Complete Battery, Terra Nova Level 10

was administered to first-grade students in the fall and the CTBS Complete Battery, Terra Nova Level 11 was administered to first graders in the spring.

Results of the difference of means test between SAGE and comparison schools on the CTBS Level 11 post-test are presented in Tables 14-17. Unlike the difference of means tests for the CTBS Level 10 pre-test, which showed no statistically significant differences between SAGE and comparison students, statistically significant differences are found in favor of SAGE students for each sub-test, and for total scale scores on the post-test.

Table 14. Differences of Means Test on Language CTBS SPRING POST-TEST for 1997-98 First-Grade Students

		SCALE SCORES		NORMAL CURVE EQUIVALENTS	
	N	**MEAN***	**STANDARD DEVIATION**	**MEAN**	**STANDARD DEVIATION**
Comparison Schools	844	573.98	46.84	50.07	21.53
SAGE Schools	1319	586.02	45.33	55.78	21.17
*Significant at .05 level					

Table 15. Differences of Means Test on Reading CTBS SPRING POST-TEST for 1997-98 First-Grade Students

		SCALE SCORES		NORMAL CURVE EQUIVALENTS	
	N	**MEAN***	**STANDARD DEVIATION**	**MEAN**	**STANDARD DEVIATION**
Comparison Schools	844	570.80	45.52	47.81	21.87
SAGE Schools	1318	580.33	41.33	52.50	20.77
*Significant at .05 level					

Table 16. Differences of Means Test on Mathematics CTBS SPRING POST-TEST for 1997-98 First-Grade Students

		SCALE SCORES		NORMAL CURVE EQUIVALENTS	
	N	**MEAN***	**STANDARD DEVIATION**	**MEAN**	**STANDARD DEVIATION**
Comparison Schools	841	525.14	42.53	45.21	19.90
SAGE Schools	1334	538.63	40.09	51.72	19.24
*Significant at .05 level					

Table 17. Difference of Means Test on Total CTBS SPRING POST-TEST for 1997-98 First-Grade Students

		SCALE SCORES		NORMAL CURVE EQUIVALENTS	
	N	**MEAN***	**STANDARD DEVIATION**	**MEAN**	**STANDARD DEVIATION**
Comparison Schools	829	556.87	38.83	47.54	21.01
SAGE Schools	1310	568.63	36.66	53.91	20.17
*Significant at .05 level					

The largest gain in SAGE student scores from pre-test to post-test, relative to comparison school students, was on the mathematics sub-test shown in Table 18. The smallest relative gain for SAGE students from pre-test to post-test was on the language arts sub-test.

Table 18. Change in Mean Score from PRE-TEST to POST-TEST for 1997-98 First-Grade Students

	Scale Scores			Normal Curve Equivalents		
	SAGE Gain	**Comparison Gain**	**Gain Difference**	**SAGE Gain**	**Comparison Gain**	**Gain Difference**
Language Arts	52.69	44.11	8.57*	10.33	6.40	3.93
Reading	45.32	34.99	10.33*	7.54	2.04	5.51
Mathematics	43.64	32.44	11.20*	7.30	1.91	5.39
Total	47.26	37.73	9.53*	9.36	4.11	5.25
*significant at .05 level						

Regression Analysis

Regression Models. The effect of the SAGE program on student achievement, controlling for other factors, was tested through a series of ordinary least squares regression models for each sub-test and for total scale scores. Control variables were entered into the models in blocks, with the SAGE/comparison student variable entered into the models last.

The first block of control variables included student score on the pre-test and school attendance, measured as number of days absent, as reported by teachers in Spring 1998. The second block of control variables included 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. A residual category, “other”, is included in the constant term in the regression equations. Eligibility for subsidized lunch, as an indicator of family income, is also

included in the second block of control variables. This variable is coded 0 if student is ineligible, 1 if student is eligible for reduced price lunch, and 2 if the student is eligible for free lunch (this variable is assumed to be interval level). In the final block, a dummy variable for SAGE or comparison school student was entered 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 19-22. For all analyses, membership in a SAGE school emerges as a significant predictor of student achievement on the post-test, while controlling for pre-test scores, family income, school attendance, and race/ethnicity. The magnitude of the effect of SAGE on student achievement, as denoted by the "b" coefficient, varies depending on the CTBS sub-test.

The largest effects of SAGE are found on the on the language sub-test, while the smallest effects of SAGE are found on the reading sub-test. When all cases are analyzed the goodness-of-fit of the models (as denoted by the adjusted R square statistic), ranges from .270 (reading sub-scale score) to .550 (total scale score). This means that when predicting the reading score and total score, the variables included in the model explain 27% and 55% of the variance respectively. Most of the variance in the post-test scores is, of course, explained by the pre-test scores.

Explained Variance in Achievement Scores. Attendance (as represented by "days absent") emerges as a consistent and statistically significant predictor of performance on all sub-tests and total scale score. "Family Income" and "Race" show some relatively large effects (as denoted by the b coefficients), but the effects are highly variable and are only sometimes statistically significant (race is discussed further below). Membership in SAGE schools has a consistently positive, statistically significant effect on achievement on the CTBS.

Table 19. SCALE SCORES Regression for Language Arts for 1997-98 First-Grade Students

	Block 1		Block 2		Block 3	
Variable	b	t	b	t	b	t
Pre-Test Score	.56	24.87*	.55	24.18*	.55	23.94*
Days Absent	-.37	-3.31*	-.35	-3.06*	-.33	-2.90*
Subsidized Lunch Eligibility	-2.63	-2.46*	-1.95	-1.72	-2.06	-1.82
African American			2.78	1.02	2.50	.92
White			5.60	2.37*	5.80	2.47*
SAGE					7.25	3.64*
Constant	292.26	23.33*	291.59	23.06*	304.36	23.28*
Adjusted R Squared	.33		.33		.34	
Standard Error of Estimate	36.49		36.45		36.30	
*significant at .05 level						

Table 20. SCALE SCORES Regression for Reading for 1997-98 First-Grade Students

	Block 1		Block 2		Block 3	
Variable	b	t	b	t	b	t
Pre-Test Score	.56	20.88*	.55	20.33*	.54	20.20*
Days Absent	-.21	-1.96*	-.18	-1.60	-.16	-1.43
Subsidized Lunch Eligibility	-3.01	-2.85*	-2.06	-1.85	-2.15	-1.94*
African American			3.07	1.15	2.80	1.05
White			7.21	3.14*	7.37	3.22*
SAGE					6.98	3.59*
Constant	285.03	19.25*	284.30	19.08*	296.11	19.48*
Adjusted R Squared	.26		.27		.27	
Standard Error of Estimate	35.73		35.64		35.50	
*significant at .05 level						

Table 21. SCALE SCORES Regression for Mathematics for 1997-98 First-Grade Students

	Block 1		Block 2		Block 3	
Variable	b	t	b	t	b	t
Pre-Test Score	.62	30.65*	.62	29.21	.61	29.13*
Days Absent	-.21	-1.96*	-.18	-1.60	-2.15	-1.94*
Subsidized Lunch Eligibility	-2.75	-3.08*	-2.44	-2.62*	-2.56	-2.76*
African American			.1.66	.74	1.44	.64
White			3.05	1.56	3.24	1.66
SAGE					7.06	4.31*
Constant	235.05	22.32*	235.69	21.88*	247.05	22.40
Adjusted R Squared	.43		.43		.44	
Standard Error of Estimate	30.05		30.05		29.88	
*significant at .05 level						

Table 22. SCALE SCORES Regression for Total for 1997-98 First-Grade Students

	Block 1		Block 2		Block 3	
Variable	b	t	b	t	b	t
Pre-Test Score	.77	38.80*	.77	37.58*	.76	37.46*
Days Absent	-.34	-3.08*	-2.44	-2.62*	-2.56	-2.76*
Subsidized Lunch Eligibility	-.63	-.86	-.53	-.68	-.62	-.81
African American			3.91	2.1*	3.72	2.01*
White			3.00	1.86	3.20	2.00*
SAGE					6.33	4.68*
Constant	167.12	15.51*	165.11	15.04*	176.57	15.80*
Adjusted R Squared	.54		.54		.55	
Standard Error of Estimate	24.53		24.51		24.34	
*significant at .05 level						

African-American Students

Among minority students in SAGE and comparison schools, African Americans clearly comprise the largest group of valid test scores – roughly 25% percent of SAGE students and 28% percent of comparison school students. In the analyses to follow, African-American students are first compared across SAGE and comparison schools on CTBS sub-test and total scale scores. Second, African-American students are compared to white students across SAGE and comparison schools on CTBS total scale scores.

SAGE vs. Comparison. Table 23 provides comparisons of means on the CTBS post-test, and change scores from pre-test to post-test. On the post-test, African-American SAGE students scored higher than African-American comparison school students on every sub-test and on total scale score. The differences between SAGE and comparison schools on post-test scores are all statistically significant. In addition, the differences between SAGE and comparison schools on mean change scores from pre-test to post-test scores are statistically significant. In other words, African-American SAGE students scored lower on the CTBS pre-test than African-American comparison school students, but made significantly larger gains than comparison school students from pre- to post-test, and surpassed African-American comparison school students on the post-test.

Table 23. African American Post-Test and Change Scores, by SAGE or Comparison School for 1997-98 First-Grade Students

SCALE SCORE	SAGE	COMPARISON	F
Language Arts			
Mean Post-Test	572.80	558.32	11.59*
Mean Change Pre to Post	56.05	38.27	16.41*
Reading			
Mean Post-Test	573.82	554.11	25.31*
Mean Change Pre to Post	50.55	25.79	31.67*
Mathematics			
Mean Post-Test	522.01	506.22	20.74*
Mean Change Pre to Post	49.06	27.50	41.99*
Total			
Mean Post Test	556.72	539.73	25.48*
Mean Change, Pre to Post	52.15	32.78	43.51*
*significant at .05 level			

African-American Males. Concern over the minority achievement gap on standardized tests has occasionally focused on African-American male students. Table 24 further distinguishes African-American SAGE and comparison school students by gender. The 1996-97 results showed that African-American male SAGE students attained comparable or higher change scores from pre-test to post-test when compared to African-American female SAGE students. The 1997-98 results show that African-American male SAGE students attained comparable or higher change scores from pre-test to post-test on the language arts sub-test, the mathematics sub-test, and the total score. However, none of these results is statistically significant.

Table 24. African-American Post-Test and Change Scores by Gender

	COMPARISON		SAGE	
	Male	Female	Male	Female
Language Arts				
Mean Post-Test Scale Score	554.74	562.18	570.99	574.27
Mean Change Pre to Post	33.96	42.69	58.64	53.65
Reading				
Mean Post-Test Scale Score	556.06	552.25	570.40	576.77
Mean Change Pre to Post	35.75	15.19	50.37	50.49
Mathematics				
Mean Post-Test Scale Score	511.48	501.24	522.82	520.93
Mean Change Pre to Post	28.90	26.08	53.37	45.01
Total				
Mean Post-Test Scale Score	540.94	538.70	554.85	558.23
Mean Change Pre to Post	34.78	30.65	53.47	50.93
*significant at .05 level				

African-American and White Achievement. African-American students scored significantly lower than white students on the CTBS pre-test total scale score, as shown in table 25. This result holds for both SAGE and comparison schools, though the gap between African Americans and whites is larger in SAGE schools. When all cases are analyzed, African-American SAGE students achieved greater gains on the CTBS total scale score than white SAGE students from pre- to post-test, closing the achievement gap (though the gap remains statistically significant). In contrast, African Americans in comparison schools achieved lesser gains and in the comparison schools the achievement gap widened.

Table 25. African-American and White Achievement in SAGE and Comparison Schools on Total Scale Scores for 1997-98 First-Grade Students

	PRE-TEST	POST-TEST	CHANGE
SAGE SCHOOLS			
African American	502.79	556.72	52.15
White	531.38	579.94	45.99
F	170.61*	96.09*	10.50*
COMPARISON SCHOOLS			
African American	510.07	539.73	32.78
White	528.60	569.02	41.14
F	52.21	90.15*	10.72*
*significant at .05 level			

Hierarchical Linear Modeling

Explanation. 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). At the origin of this problem is the fact that different classrooms often contain different numbers of students, thus those classrooms that contain greater numbers of

students have greater influence over the results of analyses done at the individual level. An analytical approach known as "hierarchical linear modeling" (Bryk & Raudenbush, 1992) was specifically designed to accommodate these types of data structures. Essentially hierarchical linear modeling (HLM) estimates individual effects by analyzing data within each class and then provides a weighted average of these effects. The effects of the class are then estimated as if all classes contained the same number of students. 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.

HLM Analyses. 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 scores and socioeconomic status (SES) measured as eligibility for subsidized lunch. The post-test scores were adjusted for these two variables at the individual level, therefore the effects may be thought of as being statistically independent of the effects of these variables. A number of different level-2 models, each containing different level-2 variables, was specified for each variable of interest. It is important to note that the "class size" variable used in these analyses measures the student-teacher ratio.

HLM Results. Table 26 provides 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 be interpreted as the classroom effects of the level-2 variables. Level-1 coefficients may be thought of as the average effect of the modeling variable on the criterion score at the individual level. The level-1 results indicate that lower 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. For example, in the Model A total score, an increase of one student in class size resulted in a drop of .828 points for the class average. Likewise, SAGE participation resulted in a 8.909 point gain in the class average on total score for Model B. A discussion of each model follows:

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. Depending on the test, an increase in class size of one person can be expected to produce a .29 to 1.12 loss in average post-test performance. The results for all scores show this effect to be significant.

Model B. *SAGE.* These models examined the effect of SAGE participation on the adjusted criterion score. Participation in SAGE shows statistically significant class average increases in all post-test scores as well. These score increases range from 7 points (reading) to 13 points (mathematics).

Model C. *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. The results show 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 D. *Class SES, Class Size.* These models examined the effect of class size on the adjusted criterion score after the classrooms were SES adjusted, viewed as the effect of class size 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 class size. The results indicate that class SES has a significant effect on the class average post-test performance. The effect of a 1 point class average gain in SES equates to between a 10 point and 13 point gain on the average post-test score, depending on the test. SES was measured on a three-point family income scale, thus a one point difference on average would be quite pronounced. Class size still has a significant effect on the post-test scores once SES has been accounted for.

Model E. *Class SES, SAGE*. These models examined the effect of SAGE participation on the adjusted criterion score after the classrooms were SES adjusted; viewed as the effect of SAGE participation once the effects of classroom SES are removed. When class SES and SAGE participation are entered in the same level-2 model, class SES has a significant effect on class average post-test performance. In addition, SAGE has a significant effect on class average post-test performance. In other words, the effects of SAGE participation on class average post-test scores, beyond those produced by SES differences, are significant on all post-test scores. In general, these effects are roughly the same as when SAGE is the only variable in the model (see model B), suggesting that SAGE classrooms and control classrooms are about equal on class SES.

Model F. *Class SES, Class Size, SAGE*. These models examined the effect of SAGE participation on the adjusted criterion score after the classrooms were adjusted for class size and SES; viewed as the effect of SAGE participation beyond the class size and SES effects. This model combines SES, SAGE participation, and class size in a single analysis. For all sub-tests, class SES once again has a significant effect on the class average post-test score. Class size has

no significant effect on the class average post-test score. Finally, SAGE had significant effects only on the mathematics sub-test.

Table 26. HLM Results for 1997-98 First-Grade Students

Source	Total	Reading	Language Arts	Mathematics
Level 1				
Pre-Test	0.870	0.627	0.625	0.712
SES	-0.784	-3.733	-1.612	-3.202
Level 2				
A. Class Size	-0.828*	-0.289*	-0.899*	-1.115*
B. SAGE	8.909*	7.009*	10.148*	13.090*
C. Class Size	-0.647	-0.734	-0.639	-0.722
SAGE	2.990	0.195	4.228	6.409
D. Class SES	-12.959*	-10.410*	-12.971*	-13.389*
Class Size	0.599*	-0.574*	-0.698*	0.883*
E. Class SES	-14.707*	-12.215*	-15.201*	-16.298*
SAGE	9.354*	7.320*	10.661*	13.428*
F. Class SES	-14.883*	-11.446*	-15.168*	-16.211*
Class Size	-0.015	-.252	-0.011	-0.027
SAGE	9.074	4.957	10.556	13.172*
*significant at .05 level				

Second-Grade Results 1997-98

Descriptive Statistics

Valid Test Scores. Analyses were conducted to assess the impact of SAGE on the 1997-98 second-grade CTBS Complete Battery, Terra Nova Level 13 post-test results. There were 1702 persisting students (i.e., students present in both the 1996-97 SAGE and comparison first-grade classrooms and in the 1997-98 SAGE and comparison second-grade classrooms), while there were 482 new second-grade students (students who were not in the program last year). However, second-grade post-test results are compared to the first-grade pre-test, as well as first grade post-test. Therefore, only those students who took both the first-grade pre-test and post-test, as well as the second-grade post-test, were used in the 1997-98 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. The number of valid test scores for the Fal1996 first-grade pre-test, the Spring 1997 first-grade post-test, and the Spring 1998 second-

grade post-test are presented in Table 27.

Table 27. Number of 1997-98 Second-Grade Students with Valid Test Scores

	SAGE	Comparison	Total
Fall 1996 First-Grade Pre-test			
Reading	1033	562	1595
Language Arts	1033	562	1595
Mathematics	1020	559	1579
Total	1008	448	1456
Spring 1997 First-Grade Post-Test			
Reading	1011	545	1556
Language Arts	1011	545	1556
Mathematics	1007	538	1545
Total	1001	534	1535
Spring 1997 Second Grade			
Reading	1037	561	1598
Language Arts	1037	562	1599
Mathematics	1043	559	1602
Total	1033	549	1582

Pre-Test (Baseline) Results. Both the first-grade pre-test and the first-grade post-test served as a baseline. Table 28 provides descriptive statistics on the scale scores from the first-grade pre-test as well as the first-grade post-test.

Table 28. Descriptive Statistics on CTBS First-Grade Pre-Test and Post-Test (SAGE and Comparison)

	FIRST-GRADE PRE-TEST				FIRST-GRADE POST-TEST			
	SCALE SCORES		NORMAL CURVE EQUIVALENT		SCALE SCORES		NORMAL CURVE EQUIVALENT	
Reading	535.20	36.83	45.31	19.78	584.17	35.65	54.44	18.50
Language Arts	532.70	42.03	45.03	20.74	583.05	37.78	54.53	17.89
Mathematics	494.55	38.27	44.11	18.05	546.59	41.56	55.88	20.41
Total	521.03	33.34	44.33	18.28	571.43	31.94	55.65	17.81

Difference of Means Test. The results from the difference of means tests between SAGE and comparison student scale scores from the Fall 1996 first-grade pre-test and Spring 1997 first-grade post-test are reported in Tables 29-32. The differences between SAGE schools and

comparison schools on the first-grade pre-test are not found to be statistically significant at the .05 level. Therefore, any differences between the first-grade pre-test and the second-grade test can be more confidently attributed to the student-teacher ratio of 15:1 in the SAGE classrooms. The differences between SAGE schools and comparison schools on the first-grade post-test are found to be significant on the total score and on all sub-scores. Therefore, any conclusions discussed regarding second-grade results must take into account the effects of the SAGE program while these students were in first grade.

Table 29. Differences of Means Test on First-Grade Pre-Test and Post-Test: Language Arts Scale Scores

	FIRST-GRADE PRE-TEST			**FIRST-GRADE POST-TEST**		
	N	**MEAN**	**STANDARD DEVIATION**	**N**	**MEAN**	**STANDARD DEVIATION**
Comparison Schools	562	530.69	43.09	545	579.01	39.75
SAGE Schools	1033	533.80	41.42	1011	586.07*	36.06
*significant at .05 level						

Table 30. Differences of Means Test on First-Grade Pre-Test and Post-Test: Reading Scale Scores

	FIRST-GRADE PRE-TEST			**FIRST-GRADE POST-TEST**		
	N	**MEAN**	**STANDARD DEVIATION**	**N**	**MEAN**	**STANDARD DEVIATION**
Comparison Schools	562	534.62	38.77	545	582.01	36.50
SAGE Schools	1033	535.52	35.75	1011	586.07*	36.06
*significant at .05 level						

Table 31. Differences of Means Test on First-Grade Pre-Test and Post-Test: Mathematics Scale Scores

	FIRST-GRADE PRE-TEST			**FIRST-GRADE POST-TEST**		
	N	**MEAN**	**STANDARD DEVIATION**	**N**	**MEAN**	**STANDARD DEVIATION**
Comparison Schools	559	493.70	38.26	538	541.88	40.75
SAGE Schools	1020	495.01	38.29	1007	550.67*	41.17
*significant at .05 level						

Table 32. Differences of Means Test on First-Grade Pre-Test and Post-Test: Total Scale Score

	FIRST-GRADE PRE-TEST			FIRST-GRADE POST-TEST		
	N	**MEAN**	**STANDARD DEVIATION**	**N**	**MEAN**	**STANDARD DEVIATION**
Comparison Schools	548	519.96	33.59	534	567.64	32.29
SAGE Schools	1008	521.61	33.20	1001	574.72*	30.97
*significant at .05 level						

As noted above, student populations varied in SAGE and comparison schools due to withdrawals and within-year enrollments. The post-test results are based only on those second graders who remained in class the entire 1996-97 first grade and 1997-98 second grade school years.

Results of the difference of means test between SAGE and comparison schools on the second-grade post-test can be seen in Table 33. Table 34 shows that when the first-grade pre-test is used as the baseline score, significant results are found on the language arts sub-scale, mathematics sub-scale, and total score. However, when the first-grade post-test is used as the baseline score, no significant results are found. This suggests that the statistically significant positive effects of SAGE occurred in the first grade. These positive effects were maintained, but did not significantly increase in second grade.

Table 33. Difference of Means Test – Second-Grade Scale Scores

	SAGE Schools			Comparison Schools		
	N	**MEAN**	**STANDARD DEVIATION**	**N**	**MEAN**	**STANDARD DEVIATION**
Language Arts	1037	610.91*	41.10	562	602.70	41.38
Reading	1037	608.17	36.11	561	604.63	37.07
Mathematics	1043	572.11*	41.69	559	564.36	39.10
Total	1033	597.14*	34.29	549	591.25	34.10
*significant at .05 level						

The largest gain in SAGE student scores from first-grade pre-test to the second-grade post-test was on the mathematics sub-test, as shown in Table 34. The smallest relative gain for

SAGE students from pre-test to post-test was on the reading sub-scale; this gain was not statistically significant.

Table 34. SAGE and Comparison Gain

	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**
Language Arts	77.07	71.74	5.33*	25.67	22.76	2.91
Reading	72.78	69.62	3.16	22.33	22.01	3.32
Mathematics	77.54	70.07	7.47*	22.86	21.97	0.89
Total	75.90	70.80	5.1*	23.67	22.36	1.31
*significant at .05 level						

Regression Analysis

Regression Models. The effect of the SAGE program on student achievement for second 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, 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 score on the first-grade pre-test or post-test and eligibility for subsidized lunch as an indicator of family income. Because attendance data were not reliably reported by districts for second graders during 1997-98, attendance was not included in the analytical model for second graders. As with the first graders (discussed earlier), the second block of control variables included dummy variables for race/ethnicity. Finally, a dummy variable for SAGE or comparison school student was entered on the third block. As with the first graders, 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 35-42. When either the first-grade pre-test or the first-grade post-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 except reading. The magnitude of the effect of SAGE on student achievement, as denoted by the "b" 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. When all cases are analyzed, the goodness-of-fit of the models (as denoted by the adjusted R square statistic), ranges from .19 to .47. Most of the variance, as was the case with the first graders, is explained by the baseline scores (either the first-grade pre-test or the first-grade post-test). "Family Income" and "Race" show some relatively large effects (as denoted by the b coefficients) and these effects are usually statistically significant. Most importantly, membership in SAGE schools has a consistently statistically significant positive effect on the language arts sub-test, the mathematics sub-test, and the total score.

Table 35. SCALE SCORES OLS Regression for Language Arts: Pre-Test as Control

	Block 1		**Block 2**		**Block 3**	
Variable	b	t	b	t	b	t
Pre-Test Score	.45	20.24*	.43	19.22*	.43	19.18*
Subsidized Lunch Eligibility	-3.82	-3.71*	-2.22	-2.09*	-2.25	-2.12*
African American			14.58	5.60*	14.22	5.47*
White			2.48	1.07	2.55	1.11
SAGE					6.44	3.46*
Constant	370.29	30.15*	382.62	31.39*	392.46	31.46*
Adjusted R Squared	.24		.26		.27	
Standard Error of Estimate	36.03		35.43		35.31	
*significant at .05 level						

Table 36. SCALE SCORES OLS Regression for Language Arts: Post-Test as Control

	Block 1		**Block 2**		**Block 3**	
Variable	b	t	b	t	b	t
Post-Test Score	.49	20.08*	.46	18.62*	.45	18.36*
Subsidized Lunch Eligibility	-4.63	-4.60*	-3.08	-2.96*	-3.14	-3.02*
African American			12.97	5.12*	12.60	4.99*
White			3.18	1.41	3.27	1.46
SAGE					6.18	3.36*
Constant	328.73	22.72*	347.62	23.92*	359.60	24.10*
Adjusted R Squared	.23		.25		.25	
Standard Error of Estimate	36.29		35.77		35.66	
*significant at .05 level						

Table 37. SCALE SCORES OLS Regression for Reading: Pre-Test as Control

	Block 1		Block 2		Block 3	
Variable	b	t	b	t	b	t
Pre-Test Score	.37	15.98*	.35	14.77*	.35	14.77*
Subsidized Lunch Eligibility	-6.46	-6.89*	-5.14	-5.27*	-5.14	-5.28*
African American			6.02	2.50*	5.84	2.42*
White			5.19	2.44*	5.22	2.45*
SAGE					3.18	1.85
Constant	416.18	32.96*	426.13	33.47*	430.56	33.26*
Adjusted R Squared	.19		.20		.20	
Standard Error of Estimate	32.93		32.68		32.65	
*significant at .05 level						

Table 38. SCALE SCORES OLS Regression for Reading: Post-Test as Control

	Block 1		Block 2		Block 3	
Variable	b	t	b	t	b	t
Post-Test	.46	20.98*	.44	19.95*	.44	19.85*
Subsidized Lunch Eligibility	-6.07	-7.06*	-4.55	-5.08*	-4.57	-5.10*
African American			5.46	2.51*	5.31	2.43*
White			6.23	3.23*	6.26	3.25*
SAGE					2.42	1.53
Constant	344.12	26.28*	352.98	26.84*	357.25	26.58*
Adjusted R Squared	.26		.27		.27	
Standard Error of Estimate	31.08		30.76		30.75	
*significant at .05 level						

Table 39. SCALE SCORES OLS Regression for Mathematics: Pre-Test as Control

	Block 1		Block 2		Block 3	
Variable	b	t	b	t	b	t
Pre-Test Score	.63	-2.80*	.58	25.04*	.58	25.23*
Subsidized Lunch Eligibility	-2.66	-2.80*	-1.21	-1.26	-1.22	-1.27
African American			17.55	7.35*	17.34	7.32*
White			1.50	.70	1.45	.68
SAGE					7.94	4.73*
Constant	262.50	22.76*	288.48	24.88*	299.02	25.49*
Adjusted R Squared	.36		.39		.40	
Standard Error of Estimate	32.80		31.96		31.74	
*significant at .05 level						

Table 40. SCALE SCORES OLS Regression for Mathematics: Post-Test as Control

	Block 1		Block 2		Block 3	
Variable	b	t	b	t	b	t
Post-Test Score	.55	28.17*	.50	25.38*	.50	25.14*
Subsidized Lunch Eligibility	-4.78	-5.36*	-3.01	-3.30*	-3.05	-3.35*
African American			15.14	6.73*	14.90	6.63*
White			4.10	2.06*	4.20	2.12*
SAGE					3.97	2.44*
Constant	274.87	25.18*	300.38	27.30*	307.76	27.01*
Adjusted R Squared	.36		.39		.39	
Standard Error of Estimate	32.22		31.41		31.37	
*significant at .05 level						

Table 41. SCALE SCORES OLS Regression for Total: Pre-Test as Control

	Block 1		Block 2		Block 3	
Variable	b	t	b	t	b	t
Pre-Test Score	.65	30.92*	.62	28.85*	.62	28.93*
Subsidized Lunch Eligibility	-2.63	-3.42*	-1.82	-2.32*	-1.84	-2.36*
African American			12.04	6.19*	11.86	6.12*
White			.01	.01	.04	.02
SAGE					5.21	3.80*
Constant	258.65	22.99*	276.79	24.39*	284.31	24.78*
Adjusted R Squared	.42		.44		.45	
Standard Error of Estimate	26.10		25.66		25.55	
*significant at .05 level						

Table 42. SCALE SCORES OLS Regression for Total: Post-Test as Control

	Block 1		Block 2		Block 3	
Variable	b	t	b	t	b	t
Post-Test Score	.69	33.44*	.65	31.12*	.65	30.88*
Subsidized Lunch Eligibility	-3.31*	-4.60*	-2.29	-3.11*	-2.33	-3.16*
African American			9.75	5.42*	9.58	5.32*
White			2.40	1.51	2.47	1.56
SAGE					2.95	2.26*
Constant	204	16.93*	225.24	18.47*	231.32	18.54*
Adjusted R Squared	.45		.46		.47	
Standard Error of Estimate	25.31		24.91		24.87	
*significant at .05 level						

African-American Students

Like the first graders, African-American second-grade students definitely comprise the largest subgroup of valid test scores – roughly 21% of all SAGE students and 25% 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.

SAGE vs. Comparison. Table 43 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 higher than comparison school students on every sub-test and on the total scale score. However, the differences between SAGE and comparison students on the second-grade test scores are not statistically significant. When using the first-grade pre-test as the baseline score, statistically significant change scores are found on all scores except for reading. However, using the first-grade post-test as the baseline score shows no statistically significant differences between SAGE and comparison schools.

Table 43. African-American Post-Test and Change Scores, by SAGE or Comparison

SCORE	SAGE	COMPARISON
Language Arts		
Mean Second-Grade Score	591.09	587.19
Mean Change From First-Grade Pre-Test to Second Grade	72.41*	62.26
Mean Change From First-Grade Post-Test to Second Grade	19.41	20.62
Reading		
Mean Second-Grade Score	594.54	543.48
Mean Change From First-Grade Pre-Test to Second Grade	73.07	68.48
Mean Change From First-Grade Post-Test to Second Grade	19.30	20.91
Mathematics		
Mean Second-Grade Score	545.44	543.94
Mean Change From First-Grade Pre-Test to Second Grade	72.59*	61.08
Mean Change From First-Grade Post-Test to Second Grade	16.46	20.58
Total		
Mean Second-Grade Score	577.04	575.94
Mean Change From First-Grade Pre-Test to Second Grade	72.44*	65.44
Mean Change From First-Grade Post-Test to Second Grade	17.99	21.14
*significant at .05 level		

African-American Males. Table 44 further distinguishes African-American SAGE and comparison school students by gender. These results show no significant gender-related gain differences when using either the first-grade pre-test or the first-grade post-test as a baseline measure.

Table 44. African-American Post-Test and Change Scores by Gender

	COMPARISON		SAGE	
	Male	Female	Male	Female
Language Arts				
Mean Post-Test Scale Score	579.63	591.49	590.74	592.22
Change (Pre-Test as Baseline)	62.03	62.52	73.74	72.09
Change (Post-Test as Baseline)	23.97	18.31	19.48	19.35
Reading				
Mean Post-Test Scale Score	584.84	596.65	590.50	598.16
Change (Pre-Test as Baseline)	66.23	69.74	71.13	74.34
Change (Post-Test as Baseline)	25.24	17.92	13.62	24.21
Mathematics				
Mean Post-Test Scale Score	538.82	547.33	550.29	540.34
Change (Pre-Test as Baseline)	55.68	65.23	74.99	69.35
Change (Post-Test as Baseline)	21.91	25.46	21.65	30.79
Total				
Mean Post-Test Scale Score	568.54	580.32	577.36	576.87
Change (Pre-Test as Baseline)	62.09	67.74	73.17	72.11
Change (Post-Test as Baseline)	22.24	20.36	16.60	19.21
*significant at .05 level				

African-American and White Achievement. African-American students scored lower than white students on the first-grade pre-test total scale score, as shown in table 45. This result is statistically significant for both SAGE and comparison schools, though the gap between African Americans and whites is larger in the SAGE schools. The change from first-grade post-test to the second-grade test shows that the SAGE African Americans kept pace with white students, but did not further close the achievement gap in second grade.

Table 45. African-American versus White Achievement on Total Scale

	First-Grade Pre-Test	First-Grade Post-Test	Second Grade	Change From Pre to Second	Change From Post to Second
SAGE					
African American	504.09	559.02	577.04	72.44	17.99
White	533.21	582.56	606.81	73.61	24.50
F	142.06*	94.70*	132.69*	.291	8.46*
Comparison					
African American	510.22	553.48	575.94	65.44	21.14
White	528.79	578.28	602.67	73.60	23.80
F	32.24*	61.80*	66.28*	7.46*	.925
*significant at .05 level					

Hierarchical Linear Modeling

Hierarchical linear models were used to evaluate the second-grade achievement results using the same series of models used to assess first-grade results. Two sets of analyses were done. The first, shown in Table 45 used first-grade pre-test as the initial achievement level of the students. The second, shown in Table 46, used first-grade post-test as the initial achievement level of the students.

HLM Results. Tables 46 and 47 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 be interpreted as the classroom effects of the level-2 variables. Level-1 coefficients may be thought of as the average effect of the modeling variable on the criterion score at the individual level. The level-1 results indicate that lower 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. For example, in the Model A total score, an increase of one student in class size resulted in a drop of .764 points for the class average. Likewise, SAGE participation resulted in a 6.322 point gain in the class average on total score for Model B. A discussion of each model follows.

Model A. Depending on the test, an increase in class size of one person can be expected to produce a .66 to 1.01 loss in average post-test performance. The results for all scores show this effect to be significant. It is noteworthy that the results of Tables 41 and 42 for the Model A are similar, indicating that the second-grade class size has little effect on the two year gain (first grade and second grade) versus the one year gain (second grade only).

Model B. Participation in SAGE shows no statistically significant class average increases in all post-test scores with the exception of mathematics in the pre-test condition. A

comparison of these results with the first-grade only results indicates second-grade SAGE effects of lesser magnitude than the first-grade effects. These results tend to support the hypothesis that SAGE is not continuing to produce the same advantage through second grade. The advantage of SAGE students is, however, maintained.

Model C. Combining class size and SAGE participation in a single analysis isolates the effects that SAGE might have beyond those produced by lower class size. The results show that once class size has been accounted for, SAGE has no significant effect on class average performance.

Model D. 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 class size. The results indicate that class SES has a significant effect on the class average post-test performance. The effect of a 1 point class average gain in SES equates to between a 10 point and 13 point gain on the average post-test score, depending on the test. SES was measured on a three-point family income scale; thus a one point difference on average would be quite pronounced. Class size still has a significant effect on the post-test scores once SES has been accounted for.

Model E. When class SES and SAGE participation are entered in the same level-2 model, class SES has a significant effect on class average post-test performance. However, for the most part, SAGE has no significant effect on class average post-test performance once class SES is controlled. In other words, the effects of SAGE participation on class average post-test scores, beyond those produced by SES differences, are not significant on post-test scores. The single exception is for the mathematics score across the two-year time span. In general, the

SAGE effects are roughly the same as when SAGE is the only variable in the model (see Model B), suggesting that SAGE classrooms and comparison classrooms are about equal on class SES.

Model F. This model combines SES, SAGE participation, and class size in a single analysis. For most sub-tests, class SES once again has a significant effect on the class average post-test score. Again, class size has a significant effect on the class average post-test score for all scores except language. Finally, SAGE has no significant effect on any sub-test, once class size is accounted for.

Table 46. HLM Results for 1997-98 Second-Grade Students – Pre-Test as Initial Achievement

Source	Total	Reading	Language Arts	Mathematics
Level 1				
Pre-Test	0.703	0.440	0.509	.662
SES	-1.929	-4.503	-1.069	-3.196
Level 2				
A. Class Size	-0.764*	-0.630*	-0.738*	-1.007*
B. SAGE	6.322*	3.302*	7.354*	8.612*
C. Class Size	-0.840*	-0.886*	-0.703	-1.073*
SAGE	1.533	5.152	0.731	1.372
D. Class SES	-9.129*	-12.214*	-7.932	-6.661
Class Size	0.678*	-0.523*	-0.666*	0.946*
E. Class SES	-10.232*	-13.214*	-8.944*	-8.251*
SAGE	5.939	3.024	7.042	8.392*
F. Class SES	-9.086*	-12.018*	-8.017	-6.613
Class Size	-0.714*	-0.715*	-0.593	-0.979*
SAGE	0.708	3.825	-1.487	0.681
*significant at .05 level				

Table 47. HLM Results for 1997-98 Second-Grade Students – Post Test as Initial Achievement

Source	Total	Reading	Language Arts	Mathematics
Level 1				
Pre-Test	0.750	0.506	0.551	.574
SES	-2.742	-4.142	-4.640	-4.703
Level 2				
A. Class Size	-0.811*	-0.669*	-0.732*	-0.959*
B. SAGE	6.296	3.651	7.432	7.741
C. Class Size	-0.916*	-0.916*	-0.683	-1.038*
SAGE	2.177	5.043	1.004	1.622
D. Class SES	-9.106*	-12.888*	-9.069*	-5.906*
Class Size	0.726*	-0.553*	-0.646*	0.911*
E. Class SES	-10.339*	-13.836*	-10.079*	-7.254
SAGE	6.007	3.300	7.040	7.616
F. Class SES	-9.023*	-12.637*	-9.172*	-5.835*
Class Size	-0.787*	-.739*	-0.556	-0.962*
SAGE	1.243	3.737	1.825	1.029
*significant at .05 level				

Effects Within SAGE Classrooms

It is of some interest to examine achievement gains in light of various classroom and teacher characteristics, teacher behaviors and student behaviors. Except for the "Teacher Experience and Achievement" analyses, data for these questions were collected only for SAGE first-grade classrooms.

Teacher Experience and Achievement

Student gains (at the classroom level) were correlated with teacher experience. In those classrooms with more than one teacher, teacher experience was averaged to obtain an appropriate experience variable for that classroom. Correlations between teacher experience and achievement gains ranged from .019 to .140, but none were significant for either the first-grade cohort or the second-grade cohort.

Student Participation and Achievement

The student participation questionnaire factored into two scales as was noted earlier. These scales were "active learning" and "on-task behavior." Interest here centered on the relationship between these variables and the classroom characteristics of class size and teacher experience as well as achievement gains. Composite active learning and on-task variables were formed for each classroom by averaging over the administrations and across students in each class. (Note, there were no significant changes in either of these variables across time). These variables were then correlated with class size, teacher experience and achievement gain in each of the three achievement sub-score areas. The on-task variable showed no significant relationships with any of the classroom variables. For the 1996-97 first-grade cohort, active learning was significantly correlated with class size ($r=-.204$; $p<.05$). For the 1997-98 first-grade cohort, active learning correlated significantly with teacher experience ($r=.307$; $p<.05$). Since these correlations did not replicate across cohorts, little practical significance should be placed in

these findings.

Proximity to the Curriculum and Achievement

From the teacher questionnaire, variables associated with proximity to the curriculum were constructed. The only significant correlation was the language arts proximity score with reading gain (r=.270) for the 1997-98 first-grade cohort. Again, since these results did not replicate across cohorts, little practical significance should be placed in this finding.

Class Organization and Achievement

SAGE students are organized into different types of classrooms as discussed earlier in the report. These classrooms include 15:1 Regular classrooms, 30:2 2-Team Teaching classrooms, 15:1 Shared-Space classrooms, and 30:2 Floating Teacher classrooms. Available data consisted of fifty-nine 15:1 classrooms, and thirty-one 30:2 classrooms. Neither the 15:1 Shared Space nor the 30:2 Floating Teacher class organization contained enough classrooms to analyze each of these four types separately. This analysis uses Hierarchical Linear Modeling to make this comparison. The HLM results show no statistically significant relationship between type of classroom organization and achievement on any of the sub-tests or the total score.

ANALYSES OF SAGE TEACHERS, CLASSROOMS, AND SCHOOLS 1997-98

Teacher and Classroom Analyses

The thesis of the SAGE program is that reduced class size, rigorous curriculum, staff development, and lighted schoolhouse activities can increase student academic achievement. However, these elements, with the possible exception of lighted schoolhouse activities, cannot influence academic achievement directly. They are mediated by classroom events. They must first influence what teachers and students do in the classroom before they can possibly have any effect on students' learning. To fully understand achievement effects in relation to the SAGE variables, it is necessary, therefore, to examine classroom changes brought about as a result of reduced class size and the other aspects of SAGE. In this section the relationship of classroom events to reduced class size, the principal SAGE variable, is examined. Data obtained from Teacher Interviews, Classroom Observations, Teacher Logs, and Teacher Questionnaires are reported below for first and second grade only. Kindergarten data are not reported at the classroom level because of the absence of corresponding achievement data. Further, it should be noted that for a variety of reasons, completed instruments, particularly Teacher Logs and Teacher Questionnaires, were not returned by all teachers, and therefore, discrepancies may occur in reported frequencies.

Teacher Interviews

Twenty-eight of the SAGE teachers who served as the observation sample were interviewed, either individually or in teams, in Spring 1998. Of this total 17 teachers were first-grade teachers, 9 were second-grade teachers, and 2 were combined first- and second-grade teachers. In terms of SAGE classroom types, the interviews were distributed in the following way: *15:1 Regular* (one teacher teaching 15 students)—10 teachers, *15:1 Shared Space* (two teachers each with 15 students sharing a room usually divided by a wall)—2 teachers, *30:2 2-Teacher Team* (two teachers teaching 30 students)—10 teachers, and *45:3 3-Teacher Team* (three teachers teaching 45 students)—6 teachers. The interviews, which lasted from 20 minutes to over an hour, were tape recorded and transcribed. They required teachers to describe the

extent to which their teaching was affected by small class size, the extent to which they believed their students' learning had improved as a result of being in a small class, and changes they anticipated making in their teaching during year three of SAGE. (See Appendix A for the Interview Guide.) Results regarding these three areas follow.

Each of the interviewed teachers indicated that his or her teaching had changed as a result of having a small-size class. The areas mentioned most frequently were knowledge of students, discipline, instruction, individualization, and learning activities. Although all of these areas were also found to be important in the Fall 1997 teacher interviews, it is becoming clearer from the Spring 1998 interviews that the most important change that results from having fewer students is individualization.

Knowledge of Students

When there are fewer students in a class teachers develop greater knowledge and understanding of each one, they indicated. This knowledge appears to be of two kinds: personality knowledge and task-progress knowledge. Because there is more time to interact with each child the teachers come to know the total child, his or her broad strengths and weaknesses. Longer parent-teacher conferences, because fewer conferences are scheduled during conference days, further help to develop this personality knowledge. The class becomes a closely-knit group or a family, as many teachers remarked. The teacher knows the students, but students also come to know each other better and are more willing to share their thoughts and problems with the class.

Task-progress knowledge occurs because there are fewer students to monitor. Teachers stated that they are able to make contact with or get around to each child on a frequent basis to identify errors and provide direction.

Illustrative Teacher Comments

You have more time to personally get to know them. Not long ago I had a little girl whose daddy traveled. He was a trucker and he was gone quite a time and her work went down, down, down, down. And I thought, OK, something is wrong. And I was able to quickly get hold of Momma and talk to her to find out what was wrong where probably with a big group she probably would have gotten lost in

the shuffle. I was able to talk about Dad and we drew a picture of him and his truck and other stuff we would not have been able to in a large class.

I know exactly where the children are and exactly what skill basics they haven't quite mastered.

You are really able to focus on each child and their problems each day.... I know what each student is doing every day.

I think that I know them really well this year. I never had a small group before; it's just like having your own little family.

Discipline. The teachers unanimously agree that the problem of class discipline is greatly reduced if not eliminated because of the small size class. Fewer discipline problems can occur because there are fewer students to misbehave, but also the family-like atmosphere that develops contributes to a lessening of inappropriate behavior. Further, when misbehavior does occur, it is noticed immediately and can be dealt with immediately, the teachers said. In a small-size class students are more on-task, attentive, and involved.

Illustrative Teacher Comments

They don't really have an opportunity really to get out of control at all.... With a small class, first of all you are going to have less problems and secondly, you are right on top of the children because you know it's a small room.

Fewer children does mean less discipline problems. I really don't have any discipline problems in my classroom this year.

I think as a teacher I am spending a lot less time on discipline, crowd management. There's lots more time for teaching and the children have a lot more time for learning. As far as discipline problems, no, I really don't have any. Discipline has not been a problem this year.

I have an excellent group and you know I attribute a lot of it to the small class size … we are a family.

Instruction. A result of less time spent on discipline is more time spent on instruction, the teachers indicated. Some teachers said that reduced time spent on keeping student records and other "paper work" because of having fewer students also resulted in more time available for teaching.

Illustrative Teacher Comments

You get more teaching time than what I did before. When you have 25 kids at one time and have to take two that are disruptive out into the hall and talk, the rest suffer.

The only thing that really changes as far as I can see is that you can teach.... You are actually teaching more of the time than you are disciplining.

I feel like I am spending more time on instruction.

It just seems like you have so much time, and you have the same amount, but it just seems like it can go so much farther and you can cover so much more.

Individualization. Every interviewed teacher mentioned individualization repeatedly as a change they have made in their teaching. Several teachers mentioned directly, and it can be inferred on the basis of comments from the other teachers, that individualization refers to helping students acquire common content or skills. It does not refer to permitting students to pursue their own objectives.

Teachers said that because of small class size they know the strengths and weaknesses of each child. They know where each is in the learning cycle and can respond appropriately. The teacher gets around to every child to offer help in a one-to-one situation. Further, the teacher can give help instantly when the class is small. In addition to this tutoring-type of individualization, the teachers indicated that they individualize by arranging the class in small groups based on perceived learning needs of individuals much more than in large classes.

The individualization of instruction is important for all students, the teachers indicated, but they believe that it has been of special benefit to poorer or struggling students, shy students, and students with exceptional needs. This kind of attention to problems which are identified early and treated early because of reduced class size result in reduced need for remediation of instruction later, they believe.

Illustrative Teacher Comments

There is more time for me to spend with each child. I can relate to them more even though relating has always been a priority for me. I feel that I really find a problem that is happening at the time.... I can find a problem and correct it.

I would say that the most important thing is that they get individualized attention.

Another thing that is a lot easier to do is to really meet the individual needs of learners. I can really plan for each kid, those that are very needy and those that are top notch.

The thing about it is if a child is having problems you can see it right away. You can take care of it then. You don't have to wait until they turn in their papers and then go back and reteach it to them. I mean, you get around to each child. And, you know it's essential that you go around and check their work. If they're having a problem you can take care of it right then rather than have them practice the skill wrong while they finish the worksheet or whatever work they are doing at their seat. I can take care of it right then before they get a chance to practice it wrong and so correct it right away. It works a lot better for the children.

I have a lot more time for individuals and small groups of children.

Learning activities. Another impact of reduced-class size on instruction is an increase in student-centered learning activities. The interviewed teachers said that they used considerably more hands-on manipulative activities, more enrichment-type activities, more interest centers, and more cooperative groups. They used more of these activities which permit students to be more independent and self-governing because they felt that having a small class gave them confidence in their ability to maintain control in situations where students have more freedom. Some teachers said that student-centered activities were used more often with a small class because having fewer students required fewer materials and resources necessary for many of these types of activities. Teachers in team-teaching situations said they could offer more student-centered activities because while one teacher is responsible for implementing the activities, the other teacher can focus on any misbehavior that might arise.

Illustrative Teacher Comments

Having only 15 children, it lends itself to a lot more hands-on teaching, more student centered.

We do a lot of hands-on. When we talk about money we have money. Kids get to work in pairs most of the time, two kids; they get to experiment with money. With measurement they are able to get up and walk around the room and everything, with a partner, where with 30 kids in the classroom, it would be very hard for them to go around the room.

I think that this year we have done a lot more hands-on type of thing.

We can do hands-on activities because there are more adults in the room. Yesterday...we were able to do a mathematics lesson about how many objects do you think can fit in this box.

Student Learning

All of the interviewed teachers stated that their students' achievement has increased considerably as a result of being in a small-size class. They report that students are moving through content at a much faster pace than first- or second-grade students normally do. They are much farther along in textbooks, sometimes even using textbooks that are usually reserved for the next-higher grade. In addition to content coverage, the teachers also report that they are able to expand and deepen students' learning. They are now able to add breadth to the content in terms of new topics of interest to the students, including greater attention to inquiry and personal learning skills, and they are able to dwell on a topic and pursue it in depth.

The teachers remarked that although all students are benefiting because of reduced class size, including students who have learning difficulties, the students who are learning at the most rapid rate are those who were in SAGE classrooms the previous year. The teachers said that these students were instantly recognizable as soon as school began in the fall.

Illustrative Teacher Comments

We have been able to touch on a lot of areas that I wouldn't have been able to touch on in a larger class.

We can talk about more things in depth because the smaller group does not get out of hand when we are discussing different things.... We're able to get through the book quicker and faster because the kids grasp the information a little bit faster.

Academic-wise, I am farther ahead than I have ever been. Our mathematics book, we are done with it. Other years we didn't cover the last three chapters.... Now I have to go to the third-grade teachers to ask what they would like me to work more on.

I think that they are farther along than groups that I have had in the past, in their reading.

Anticipated Changes

Although teachers were asked to think about changes they planned to make in their teaching for the third year of SAGE, most took the opportunity to either express their satisfaction with the SAGE program or to identify general problems usually without offering solutions. The satisfaction with SAGE, apart from its benefit to students, was that it made teaching more pleasurable. Reduced class size results in reduced stress. One can concentrate on actual teaching rather than having to spend time on behavioral problems, excessive paper work, and other problems. The teachers who taught in 30:2 team situations saw the teaming aspect as an additional strength. Some, in fact, appear to be unable to separate SAGE from team teaching. Teaming enables teachers to specialize in terms of content areas, reduce management because one teacher is often free to monitor the class while the other teacher teaches, discuss strengths and weaknesses of students, and cooperate in other ways.

Some of the problems mentioned were that teacher inservice was needed and more hands-on activities need to be used.

Illustrative Teacher Comments

I think that it has been a very rewarding year and the children have grown tremendously. It just has been wonderful to continue SAGE into the upper grades. I think it is just an excellent, excellent program.

I have just enjoyed it so much to be honest.... To me sharing a class has really brought a new dimension to the whole thing.... That person is in there constantly and she has expectations that I might have slipped by. Or, she will catch something that could just go by the wayside because I am busy doing something else and it's really been to the benefit of the children as well as both of us, I think.

I would like to get more into, you know, some more open-ended activities, things like that. This year...we were kind of feeling our way.

The only way that SAGE will work for me is if we have kids that are in the program for the whole SAGE.... With the mobility within the city kids are moving. I lost two kids this year who were good students, and it really broke my heart to know that they went into a bigger classroom.

Classroom Observations

Observations were made in selected SAGE classrooms during the 1997–98 school year. The classrooms were chosen to represent different geographic areas of the state, grade levels, and types of SAGE classrooms. The sample consists of 25 classrooms in 12 schools from 9 school districts; 14 first grades, 9 second grades, and 2 combined grades; and fourteen 15:1 Regular, four 15:1 Shared Space, five 30:2-Teacher Team, and two 45:3-Teacher Team classrooms.

Each classroom was observed twice, once during the fall and once during the spring, with the exception of a second-grade classroom that could only be observed during the fall. The observations, which took place in either reading, language arts, or mathematics, were open-ended observations designed to capture a broad range of classroom events regarding teaching and learning. Following the observations during which observers took careful notes, expanded accounts were written for each observation. These accounts were then analyzed using a set of categories developed from observations made during the first year of SAGE. The main categories, each of which has subcategories, are individualization, engagement, and management. (See Appendix B for the Observation Guide.) Classroom behavior related to these three areas is discussed below for the total group of classrooms, for first and second grades, and for types of SAGE classrooms.

Total Classroom Behavior

Behaviors expressed as percents of total behavior in individualization, engagement, and management for fall and spring are presented in Table 48. It can be seen that, except in a few areas, the behavior observed in the fall is similar to that observed in the spring.

Table 48. Total Classroom Behavior for Fall and Spring

	FALL		SPRING	
	N of cases	**Percent**	**N of cases**	**Percent**
Individualization				
Monitoring	25	11	25	13
Grouping	25	9	25	8
Choice	25	2	25	1
Help	25	26	25	23
Participation	25	47	25	39
Whole Class	25	4	25	12
All Children	25	2	25	4
Engagement				
Listening	25	43	25	54
Practicing	25	10	25	3
Responding	25	28	25	22
Gaming	25	2	25	1
Manipulating	25	3	25	3
Creating	25	2	25	4
Dialoguing	25	4	25	4
Problem Solving	25	2	25	2
Reporting	25	2	25	4
Reflecting	25	0	25	0
Initiating	25	3	25	2
Management				
Praise	25	33	25	23
Reproof	25	13	25	7
Remind	25	23	25	16
Warms	25	4	25	4
Cools	25	1	25	2
Peer	25	8	25	18
Permits	25	19	25	31

Individualization. The classroom observations and teacher interviews yield quite consistent data in the area of individualization. Table 48 reveals that about 90% of the observed classroom time is spent in some form of individualized activity in which students are working on their own or in groups on selected or assigned tasks and being monitored or helped, or they are actively participating in a group discussion by expressing their views and understandings. For only about 10% of the time are they being instructed by the teacher as a total class. This general pattern is constant over the course of the year with two possible exceptions. Active participation is somewhat less in spring and total class instruction with a common task is somewhat more.

Engagement. For both fall and spring engagement consists more of teacher-centered instruction than student-centered instruction. Listening, practicing, responding, and other

activities directed and controlled by the teacher, make up 80% of the classroom events. The remaining 20% of engagement consist of more student-centered activities, such as manipulating, problem solving, creating, and others. On the surface, this comparatively little use of these more "hands-on" activities seems to clash with teachers' interview comments, but in the interviews teachers discussed their teaching in total while the observations only dealt with reading, language arts, and mathematics, subjects which may present fewer opportunities for student-centered activities than science and social studies.

The only change that appears to occur over the year regarding engagement is in the teacher-centered behaviors. In spring more listening but less practicing and responding occurred.

Management. Negative management comments of “reproof” and “cools” (i.e., ignoring or discouraging children) from the teachers in the observed sample are comparatively low. They make up 14% of the statements in fall and 9% in spring. Even if “reminds”, a behavior designed to preempt misbehavior, is included in negative management, negative management still totals only 37% in fall and 25% in spring. The reduction of the need for discipline in small-size classes about which the teachers spoke in the interviews is evident in these figures.

The fall-to-spring changes appear to be more marked in management than in individualization and engagement. There is less negative management in spring compared to fall, but there is also less positive management. The beneficiary of these changes is “student” and “peer” self-management. “Peer” management and “permits” management both are used more in spring.

First-Grade and Second-Grade Behavior

The classroom behavior observed in first grade and in second grade appears to be almost identical in each of the three areas as can be seen in Table 49. About 90% of the time the focus is on individuals or small groups, while 10% of the time it is on total group tasks; 80% of the engagement is more teacher centered and 20 % is more student centered; and about 30% of the management is negative, including preemptive negative, and 70% is positive, including student self-discipline.

Table 49. Classroom Behavior for First Grade and Second Grade

	GRADE 1	**GRADE 2**
Individualization	**N = 28**	**N = 16**
Monitoring	14	9
Grouping	6	14
Choice	2	0
Help	26	23
Participation	43	42
Whole Class	6	10
All Children	3	1
Engagement	**N = 28**	**N = 15**
Listening	50	45
Practicing	5	11
Responding	26	23
Gaming	1	2
Manipulating	4	2
Creating	4	1
Dialoguing	4	6
Problem Solving	2	4
Reporting	2	5
Reflecting	0	0
Initiating	3	2
Management	**N = 27**	**N = 14**
Praise	31	22
Reproof	9	14
Remind	19	21
Warms	4	3
Cools	1	2
Peer	12	11
Permits	24	26

Classroom Behavior in Different Types of SAGE Classrooms

Table 50 reports the observed classroom behavior for four types of SAGE classrooms: 15:1 Regular, 15:1 Shared Space, 30:2-Teacher Team, and 45:3-Teacher Team. Again the results are generally uniform across these four types of SAGE as they were in relation to grade level. Claims about differences in terms of type of SAGE classroom can only be speculative because the number of teachers observed in each type of classroom is small. Some of the possible differences in individualization are that 15:1 Regular teachers use large-group instruction more than the other types and 15:1 Shared Space teachers work with individuals and small groups less than the other types. In terms of engagement the types are very similar except for dialoging which 15:1 Shared Space teachers seem to employ more than the other teachers. It

is possible that constrained classroom space may cause 15:1 Shared Space teachers to keep the class together more and use discussion more because there is less space for individual and small groups to work without interruption. Management seems quite similar with the exception of 45:3-Teacher Team, and possible 30:2-Teacher Team, where more student self-directed discipline is used.

Table 50. Classroom Behavior for Different Types of SAGE Classrooms

	15:1 Reg	**15:1 SS**	**30:2 TT**	**45:3 TT**
Individualization	**N = 27**	**N = 6**	**N = 10**	**N = 4**
Monitoring	12	6	13	16
Grouping	8	9	9	12
Choice	1	0	1	4
Help	26	23	27	14
Participation	38	56	45	49
Whole Class	11	5	3	4
All Children	4	1	1	1
Engagement	**N = 27**	**N = 6**	**N = 10**	**N = 4**
Listening	48	44	49	53
Practicing	8	3	6	7
Responding	27	21	24	21
Gaming	1	0	3	2
Manipulating	3	3	4	1
Creating	3	1	6	1
Dialoguing	2	18	2	2
Problem Solving	2	2	3	6
Reporting	3	6	2	4
Reflecting	0	0	0	0
Initiating	3	2	2	4
Management	**N = 26**	**N = 5**	**N = 10**	**N = 4**
Praise	32	25	23	20
Reproof	9	20	11	6
Remind	16	28	27	17
Warms	4	2	6	5
Cools	0	1	3	0
Peer	14	12	11	8
Permits	26	12	19	45

Observation Summary

Altogether, the observations of SAGE teachers suggest that teachers in small-size classes spend the vast majority of class time actually instructing instead of managing the class. Further, the instruction is teacher controlled and directed. Rather than being primarily total class instruction, however, the instruction is largely individualized in the sense that the teacher

constantly checks with each child or a group of children during a lesson to evaluate progress and offer help, or the teacher provides opportunities for children to actively participate in class events and thereby articulate their current level of understanding which then can be corrected if needed. Student-centered activities, although present in small-size classes, are not a major feature of the classroom. The following observation example illustrates these dominant traits of small-class size teaching and learning.

Observation Example

Mrs. Donald and her class of 14 second-grade students share a classroom with Mrs. Johnson and her class of 15 first-grade students. The room is divided by a movable wall. In Mrs. Donald's side of the room, the students' desks are arranged in four rows facing the chalkboard. A calendar, class rules, U.S. map, spelling words, and other charts are fixed to the walls. The rear of the room contains Mrs. Donald's desk, two computers, and a worktable.

The lesson began with Mrs. Donald reviewing CVCV (consonant, vowel, consonant, vowel) words that the class had previously learned and on which they had been tested. She said to the class, "Who can give me an example of a CVCV word?" Various children answered, "wave," "cane," "cake," apple." "Apple? Let's put 'apple' on the board. Does 'apple' have a CVCV pattern?" she asked. The class replied in union, "No!" She continued, "What type of letter does 'apple' begin with?" Again the class responded in union, "A vowel!" Mrs. Donald then said that it is a vowel and, therefore, did not fit the CVCV pattern and that "apple" is not a good choice.

Mrs. Donald then moved on to the primary purpose of the lesson. She asked, "Does anybody remember what other types of words we talked about yesterday?" Muhammed responded, "Compound words." Mrs. Donald praised Muhammed and began a discussion of compound words by asking the class what they believe that a compound is. Ashley gave the example "cupcake," and then said, "It is two words put together to form one word." Mrs. Donald checked the students' understanding by writing "cup" and "cake" on the board and asking if what she had written was a compound word. Students replied, "No." When called on, Julie explained that separate words to not make a compound word.

Following this overview of compound words, Mrs. Donald introduced the main task of the lesson. She said, "Each of you will receive a snowman that I have already cut out of construction paper for you. You will also receive a hat with a compound word on it and a scarf. Your job is to draw the two pictures of two words that make up your compound word. Let's look at the one I have done. What is the word on the snowman's hat? Snowman, right? I drew two words together; it makes the word 'snowman.'" After responding to several children regarding questions they had about compound words, the students began the snowman activity. Mrs. Donald constantly moved about the classroom answering questions and offering help to each child. At one point Mrs. Donald stopped the class to explain that glue is not needed to attach the scarf to the snowman because the scarf is made of sticky tape. At another point, as she was circulating, Mrs. Donald said, "Gloria, you are wasting time," and "Robin, you did not do what I told you."

When students had completed the snowman activity the teachers called each to the front of the room to say their compound word and show their snowman. The compound words shared and displayed by the children included bluebird, football, strawberry, catfish, starfish, baseball, boyfriend, blueberry, bedroom, and basketball. Jordan, who was walking around the room during this sharing period, was reminded of the rules of the classroom and told to take his seat.

When Keith, the last child, had reported on his compound word, Mrs. Donald asked the class, 'What did you learn about compound words today?" She answered her own question by stating that when the students came upon a new word they should investigate the word to determine if they know any part of it. If so, it may be a compound word that they could figure out. She then introduced a new set of compound words such as "overnight" and "steamboat" and asked students to define them. Several students gave their interpretations of each of the terms.

"OK, class," Mrs. Donald next responded, "next time you see a big word you need to investigate to see if it is a compound word, then look to see if you can figure out what the words are that make up the compound word."

Mrs. Donald then transitioned into a reading lesson.

Teacher Logs

Teachers in kindergarten, first, and second grade in SAGE schools were asked to complete three teacher logs during the year. The log required teachers to record classroom events for a full day at 15-minute intervals regarding type of time, grouping, content, and student learning activities. (See Appendix C for the Teacher Log.) The numbers of logs completed and returned during each of the log data periods were the following: fall–326 (K–105, first–106, second–107), winter–296 (K–92, first–97, second–100), and spring–195 (K–92, first–97, second–99).

Log data regarding type of time, grouping, and content are reported here for first grade and second grade. Data concerning student learning activities are not reported because they were found to be unreliable. Many teachers had difficulty in completing the logs in this area, and subsequently, made numerous recording errors.

Total Log Results

Total results from the Teacher Logs concerning type of time, grouping, and content are reported in Table 51. Regarding type of time, it can be seen that almost two-thirds of classroom time is spent on actual instruction. If time spent on planning and evaluation is added to

instructional time the total time spent on education during a typical school day in a small size class averages nearly 80 percent. The remaining 20 percent is spent on classroom routines and personal time for the teacher.

Table 51. Total Log Results

ITEM	FIRST GRADE (N=285)	SECOND GRADE (N=291)	TOTAL (FIRST & SECOND)
TYPE			
Routines	9	9	9
Instructional Time	63	61	62
Planning & Evaluation	13	16	15
Personal Time	9	8	9
Housekeeping/Clerical	6	6	6
GROUPING			
Whole Group	53	51	52
Small Group	29	29	29
Individual	14	17	16
Combined Classes	3	3	3
CONTENT			
Reading/Language Arts	46	45	45
Mathematics	19	19	19
Integrated	19	19	19
Other	15	17	16

The results in relation to grouping are that about half of the instructional time is spent in whole-class instruction and slightly less than half is spent in small-group and individualized instruction. Although spending almost half of the classroom time in small groups or working with individuals suggests a very strong emphasis on individualization, this finding seems to contradict data obtained from observations which indicated that SAGE teachers spend about 90% of the classroom time in individualized instruction. Data from the two instruments are compatible, however, because teachers probably included what was meant by the observation category of student participation in their interpretation of whole-group instruction when completing the log. When appropriate parallels are made between the two instruments the two sets of data are nearly identical. They both reveal extensive use of individualization.

In terms of content, mathematics, reading, and language arts dominate instruction. They account for 65 percent of the classroom time when taught directly and another 20 percent when taught indirectly through content integration.

The data obtained from the 1997–98 logs are consistent with the 1996–97 log data.

First-Grade and Second-Grade Log Results

The results for grade one and grade two for each log period, which are also reported in Table 52, show little variation over the year and by grade level. The only slight variation that occurs regarding the three log periods is that instructional time and small-group instruction increase and planning time and individual instruction decrease at both grade levels over the year.

Table 52. First-Grade and Second-Grade Log Results

ITEM	FIRST GRADE			SECOND GRADE		
	Fall N=103	Winter N=97	Spring N=97	Fall N=79	Winter N=100	Spring N=99
TYPE						
Routines	11	8	8	10	8	8
Instructional Time	58	65	65	58	62	63
Planning & Evaluation	15	12	12	16	15	15
Personal Time	10	8	8	8	8	8
Housekeeping/Clerical	7	6	6	8	5	5
GROUPING						
Whole Group	55	52	52	52	51	51
Small Group	26	32	32	23	31	31
Individual	16	13	13	23	15	15
Combined Classes	4	3	3	3	3	3
CONTENT						
Reading/Language Arts	45	48	49	46	45	45
Mathematics	19	20	20	18	20	20
Integrated	19	18	18	19	19	19
Other	17	14	14	17	17	17

Log Results by SAGE Classroom Type

As can be seen in Table 53 little variation also exists in log data related to the different types of SAGE classrooms. The only differences that are evident are that 15:1 Shared-Space teachers report spending more time on instruction than the other teachers, 30:2-Teacher Team classes use small groups and individualized instruction more than the other classes, and 30:2 Floating Teacher classes spent more time in mathematics, reading, and language arts instruction than the other teachers.

Table 53. Log Results by Classroom Type

ITEM	15:1 REG N=273	15:1 SS N=42	30:2 TT N=237	30:2 FT N=18
TYPE				
Routines	7	11	10	13
Instructional Time	60	66	64	59
Planning & Evaluation	17	9	13	15
Personal Time	9	11	8	7
Housekeeping/Clerical	7	4	5	6
GROUPING				
Whole Group	54	49	50	52
Small Group	26	31	35	29
Individual	18	12	15	10
Combined Classes	3	8	1	9
CONTENT				
Reading/Language Arts	45	47	44	44
Mathematics	19	16	19	27
Integrated	20	19	20	15
Other	16	17	17	15

Teacher Questionnaires

Teacher questionnaires were completed by SAGE teachers during Spring 1998. The questionnaire, which elicits perceptions regarding classroom teaching, mathematics, reading, and language arts curriculum, family involvement, professional development, and overall SAGE satisfaction, was returned by 228 SAGE teachers (K– 72, first–72, second–78).

The classroom teaching section which contributes to this examination of classroom events in reduced-size class consists of 11 items. The teachers were to indicate the level of their agreement with each item and then to select and rank the three which represented the most significant ways their teaching had been affected by a reduced-size class. The results of these two analyses for the total group of first- and second-grade teachers, for first- and second-grade teachers separately, and for each type of SAGE classroom follow.

Total Questionnaire Results

Table 54 contains the results for first-grade and second-grade teachers combined. An examination of this table reveals that the teacher behaviors that received the highest ratings and rankings are more individualized instruction; more teaching time; more discussion, sharing, and answering; more hands-on activities; and more content coverage. Those teacher behaviors that

received the lowest ratings and rankings, comparatively speaking, are more integrated content; more activities based on students' prior knowledge; more use of cooperative groups; more student choice in learning activities; and content covered in more depth. These data generally confirm the results of the teacher interviews, the classroom observations, and the teacher logs. Individualization is again seen as the most important classroom product of reduced-class size. Class participation, which can be viewed as a type of individualization because individual students received answers to their questions, voice their understandings, and receive personal critique, is also an important product as is more time spent on teaching as opposed to disciplining. With the possible exception of hands-on activities, these results again suggest that the type of teaching used in small-size classes is teacher-centered, teacher-controlled teaching. Hands-on activities, although reported both here and in the interviews to be used more often because of reduced-class size, were, as has been shown, not frequently observed. As mentioned earlier, the teacher questionnaire data, like the interview data, are broader in terms of subject matter focus than the observations.

Table 54. Total Questionnaire Results (N=150)

ITEM	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Ranking Percents
1. More time teaching	1	1	7	42	49	17
2. Covered more content	0	1	5	40	54	8
3. Integrated content	0	0	12	50	39	4
4. More depth	0	1	9	46	44	5
5. Planning/ Implementing	0	0	3	30	68	25
6. More engaging	0	0	2	42	56	11
7. More Hands-on	0	0	7	56	38	10
8. Student's knowledge	0	1	15	49	36	4
9. Problem solving	1	1	6	51	42	6
10. Cooperative groups	0	4	19	42	36	5
11. More opportunities	0	4	14	47	35	5

The teacher behaviors receiving the lowest ratings and rankings again suggest that teachers are less student-centered than teacher-centered in their teaching of reduced-size classes. Student choice, independence, and interest are of less concern than individual content coverage.

Questionnaire Results for First Grade and Second Grade

The questionnaire results found for the total group of SAGE teachers are also the results found for first- and second-grade teachers separately, as can be seen in Table 55. The only slight difference that is apparent is that second-grade teachers report more content coverage as a result of small class size.

Table 55. Questionnaire Results for First Grade and Second Grade

	FIRST GRADE N=72						SECOND GRADE N=78					
ITEM	**SD**	**D**	**N**	**A**	**SA**	**%**	**SD**	**D**	**N**	**A**	**SA**	**%**
1	1	3	8	42	46	18	0	0	6	41	53	15
2	0	3	6	40	51	7	0	0	5	41	54	11
3	0	0	11	50	39	3	0	0	13	51	36	3
4	0	1	8	47	43	5	0	0	9	45	46	7
5	0	0	3	31	67	25	0	0	3	31	67	27
6	0	0	7	38	56	9	0	1	3	46	51	9
7	0	0	7	38	56	9	0	1	6	41	51	8
8	0	0	15	44	40	3	0	1	15	55	28	3
9	1	0	6	56	38	6	0	1	5	49	45	7
10	0	7	14	44	35	5	0	1	23	40	36	6
11	0	3	11	46	40	7	0	5	17	49	30	5

Questionnaire Results for Different Types of SAGE Classrooms

Generally, the results regarding classroom teaching reported by teachers in each of the types of SAGE classrooms are quite similar. Teachers from each type of SAGE gave their highest ratings to teaching time, individualization, student engagement, content coverage, and hands-on activities. There were some differences among the four types of classrooms, however. In comparison to the other teachers, the 15:1 Shared-Space teachers used hands-on activities less, the 15:1 Regular teachers used problem solving more, the 30:2-Teacher Team teachers used students' prior knowledge more, and the 30:2 Floating Teacher teachers used more time for instruction.

Table 56. Questionnaire Results for Different Types of SAGE Classrooms

	15:1 Reg Ratings (N=86)					15:1 SS Ratings (N=9)					30:2 TT Ratings (N=41)					30:2 FT Ratings (N=5)				
Item #	**SD**	**D**	**N**	**A**	**SA**	**SD**	**D**	**N**	**A**	**SA**	**SD**	**D**	**N**	**A**	**SA**	**SD**	**D**	**N**	**A**	**SA**
1	1	0	7	48	44	0	0	11	22	67	0	0	5	39	56	0	20	20	40	20
2	0	1	7	38	54	0	0	11	22	67	0	0	0	49	51	0	0	0	60	40
3	0	0	9	48	43	0	0	11	44	44	0	0	15	61	24	0	0	40	60	0
4	0	1	9	45	44	0	0	11	33	56	0	0	7	46	46	0	0	0	100	0
5	0	0	1	24	74	0	0	0	33	67	0	0	5	34	61	0	0	0	60	40
6	0	0	0	41	59	0	0	0	44	56	0	0	2	46	51	0	0	0	80	20
7	0	0	5	40	56	0	0	22	11	67	0	2	2	39	56	0	0	20	60	20
8	0	0	20	47	34	0	0	11	44	44	0	2	2	56	39	0	0	20	60	20
9	0	0	4	52	44	0	0	11	11	78	2	2	7	51	37	0	0	20	80	0
10	0	4	12	48	37	0	0	33	33	33	0	7	27	27	39	0	0	40	40	20
11	0	2	12	51	35	0	11	22	11	56	0	5	17	51	27	0	0	20	60	20

Teacher and Classroom Summary

The results from the Teacher Interview, Classroom Observation, Teacher Log, and Teacher Questionnaire support and extend those obtained in 1996–97. They demonstrate that the major change that takes place in teaching when teachers teach a reduced-size class is not a total adoption of more student-centered teaching. Teachers do not suddenly permit students to set goals or decide on learning activities, nor do they install a problem-solving approach rich with resources and manipulatives. Reduced-class size permits some movement toward more student-centered teaching, but the main effect of reduced class size appears to be a focus on students as individuals. Many, if not most, of the techniques and methods that they use may be the same techniques and methods that they have used in normal-size classrooms. The difference is that now the techniques and methods are directed at individuals much more frequently. They know each student's learning needs, they correct misunderstanding instantly, and they move ahead when the time is right. This attention to individuals is done in one-to-one situations, in small groups formed on the basis of need, and in total class situations through response and critique, and it is a continual, pervasive feature of classroom life.

What appears to be happening as teachers change from teaching large-size classes to teaching small-size classes may not be too different from what may happen to chefs as they change from cooking in a restaurant to cooking for their families. In the restaurant, the chef

prepares five or ten menu meals. Each meal generally does not differ in elements, portions, or presentation. One plate of chicken breast and pasta, for example, is identical to every other, and it is served night after night. At home the chef can pay more attention to the desires or needs of family members. He or she can vary the menu, making sure that broccoli is not put on Alice's plate, reducing the amount of salt for Grandma, and giving Joe two pieces of fish. The actual food preparation in the two settings is probably very similar, however. The chef alters his or her behavior as he or she moves from the restaurant to the home, but meal planning, food purchasing, food preparation, and food cooking do not change. The vegetables get washed in the same way, the rolls are baked in the same way, and the coffee is ground in the same way.

Although all conclusions about teaching small-size classes must be tentative at this time, a model of teaching small-size classes is beginning to emerge. This model, displayed in Figure 1, emphasizes individualization, but contains other related elements. The model speculates that having fewer students permits teachers to know them better. This knowledge aids in reducing misbehavior, which in turn makes more time available for instruction because time is not required to discipline students. More time for instruction and greater knowledge of students come together to permit more individualized instruction. More time for instruction also permits somewhat more use of hands-on activities using regular teacher-centered methods. The outcome of this heavy focus on individuals is more content coverage and, the model continues to speculate, more student achievement.

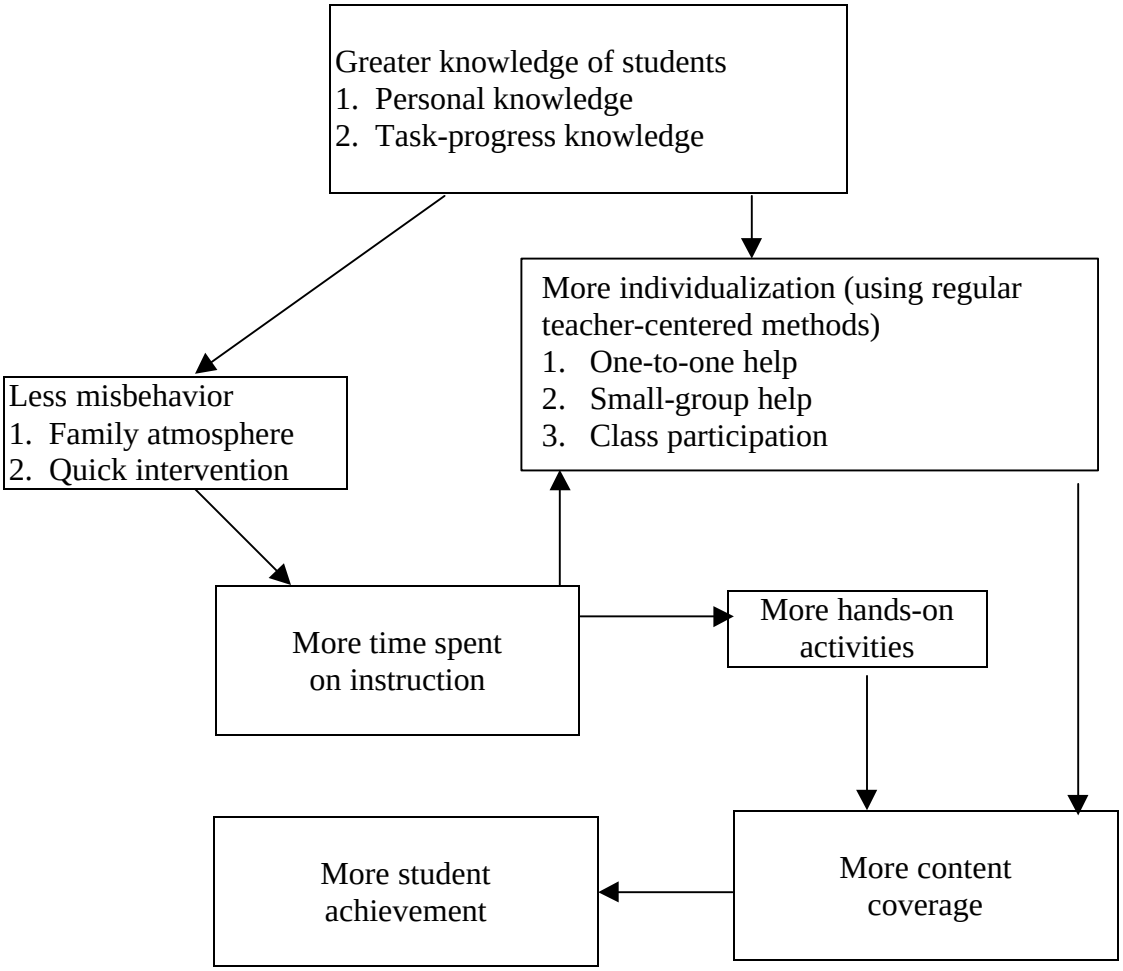
The pattern of teaching that Figure 1 depicts applies across grade levels and types of SAGE classrooms. Few differences in grade level or type of classroom were revealed in classroom data from any of the instruments. This finding is puzzling for several reasons. In terms of grade level, one might expect some differences between first and second grade because most of the first-grade teachers have experience in teaching reduced-size classes and most of the second-grade teachers do not. Further, one might also expect differences because achievement results show that although second-grade children maintained their achievement advantage over comparison school children, they did not improve on it. The same type of instruction as first

grade but less achievement than first grade could be explained, in part, on the basis of time. In at least nine SAGE schools, reduced size classes at the second-grade level were not implemented until late fall or winter.

In terms of type of SAGE classroom, differences might be expected because having ownership for 15 students as occurs in 15:1 student-teacher ratio classrooms is clearly different from two teachers sharing ownership of 30 students as occurs in team-taught classrooms. Teamed teachers said during interviews that they often use an approach in which one teacher teaches and the other monitors either offering help or disciplining if needed, but indications of this approach or other methods peculiar to team teaching were not evident in the data from the other instruments.

Results regarding classroom events from 1997-98 as well as 1996-97 suggest the need to focus future study of classroom events more specifically on the themes that have emerged. Individualization, the practice that seems to be the main effect of having a reduced-size class, needs to be examined in greater depth as do other aspects of teaching in reduced-size classes and potential variations in classroom events across grade levels and types of SAGE classrooms. The data that are now needed require case studies of selected SAGE schools and classrooms. In the 1998-99 SAGE evaluation results of the first SAGE case studies will be presented.

Figure 1. A Tentative Model of SAGE Small-Class Size Teaching and Learning



Schools

Rigorous Curriculum

The extent to which the reading, language-arts, and mathematics curricula in SAGE schools conformed to national standards is reported in Table 57. These data, which are derived from teacher perceptions from the Teacher Questionnaire, suggest general overall agreement with the standards in both curriculum areas, but somewhat greater agreement in reading and language arts than mathematics.

In reading, the areas of greatest agreement are (1) that the names of parts of books are taught, (2) students are encouraged to choose books of personal interest, and (3) students are taught to apply a variety of decoding strategies. The areas of least agreement are (1) that students are taught to critique non-print media, (2) students are taught how language can be adjusted for different audiences, and (3) students are introduced to texts representing a range of historical periods.

In mathematics, the areas of greatest agreement are (1) that students have the opportunity to connect mathematics to everyday situations, (2) students learn the enumeration system through concrete experiences, and (3) students have opportunities to deal with a wide variety of patterns. The areas of least agreement are (1) that calculators are used, (2) the concept of chance is explored through actual events, (3) perimeter and related areas are developed intuitively, and (4) metric and other nonstandard measures are taught.

The results concerning rigorous curriculum for 1997-98 are strikingly similar to those found in 1996-97. In both years, the areas that are comparatively low are areas that many would see as being inappropriate for primary school children such as use of calculators, critiquing non-print media, or learning about perimeter.

Table 57. Rigorous Curriculum, Grades K-2

Reading/Language Arts	**N**	**Mean**	**Standard Deviation**
Students introduced to texts: represent range of genres	304	4.18	.74
Students introduced to texts: represent range of historical	308	3.32	.83
Students introduced to texts: deal with topics relevant to real world	309	4.16	.69
Students introduced to texts: variety of ethnic, culture contexts	309	4.11	.72
Students taught to apply variety of decoding strategies	307	4.61	.63
Students introduced to variety of interpretative strategies	306	3.89	.91
Students taught names for parts of books	311	4.00	.98
Students introduced to literature terminology	310	4.29	.86
Students taught to categorize texts: fiction or non-fiction	311	4.78	.46
Students taught to categorize texts: topic or theme	309	3.87	.89
Students taught to categorize texts: author	309	3.83	.99
Students taught to make associations among texts	311	3.67	.78
Student taught aware of how language can be purpose adjusted	310	3.52	.85
Students taught aware of how language can be audience adjusted	309	3.26	.86
Students encouraged to choose books interested in reading	311	4.76	.46
Students apply lang/conventions: critique/discuss print texts	307	3.52	1.04
Students apply lang/conventions: critique/discuss non-print media	306	3.15	.96
Students apply lang/conventions: writing to develop interests	308	4.19	.82
Students apply lang/conventions: speaking to develop interests	309	3.98	.90
Mathematics			
Students write own mathematics problem about real or imaginary	308	3.15	.91
Students encouraged to develop own strategy for solving problems	312	3.95	.78
Opportunity to investigate open problems have more than one sol.	311	3.61	.88
Write in math class to reflect and demonstrate understanding	309	3.30	2.49
Mathematics language and symbols introduced in context of expl.	311	4.12	.72
Opportunities to make connections between mathematics and other	312	4.14	.62
Opportunities to make connections between math & everyday	205	4.27	.63
Estimation when working with quantities, measurement, comput.	312	3.72	.79
Opportunity to explore and use estimation strategies in real sit.	311	3.65	2.42
Learn enumeration through concrete experiences	311	4.23	.77
Discuss, model, draw, write about their understanding	309	3.71	.89
Instruction of facts emphasize development of thinking strategies	309	4.06	.81
Develop own computation strategies and algorithms	304	3.34	1.08
Calculators used in appropriate situations	310	2.46	1.23
Instruction includes concrete experiences with metric units	309	3.11	1.13
Concepts of perimeter, area, volume are developed	306	3.12	.93
Opportunity to explore geometric shapes through concrete exp.	310	3.87	.78
Opportunity to work with 3-dimensional figures	310	3.54	.93
Formulate & solve problems involving collecting & analyzing data	311	3.57	.79
Make predictions, inferences, decisions from data	309	3.83	.77
Concept of chance explored by collection of data and other events	308	3.10	.84
Concrete and real experience to develop fraction concepts	309	3.61	.86
Recognize, describe, extend patterns	311	4.23	.68
Create patterns using materials and discuss patterns	311	4.16	.76

Professional Development

The professional development section of the Teacher Questionnaire was completed by SAGE kindergarten, first-grade, and second-grade teachers in May 1998. A total of 150 teachers at the first- and second-grade levels returned the questionnaire. Results from these questionnaires regarding context and process of professional development programs in the teachers' schools and teachers' personal development plans are contained in Tables 58 and 59.

Table 58. Teachers' Perceptions of Professional Development (Grades 1 & 2) N=150

ITEM	STRONGLY DISAGREE	DISAGREE	NEUTRAL	AGREE	STRONGLY AGREE
1. Ongoing & Regular	1	4	7	42	46
2. changes in Practice	1	7	26	45	20
3. Adequate Funding	4	17	15	49	15
4. Widespread Support	2	12	25	43	19
5. Joint Learning	22	34	21	16	7
6. Study Groups	15	24	22	30	10
7. Improvement Plan	1	12	21	51	16
8. "Teacher as Learner"	0	6	21	57	17
9. Staff Development	1	6	16	60	17
10. Precede Decisions	3	19	43	31	6
11. Program Evaluation	3	18	41	30	8
12. Staff Development Activities	3	25	33	33	7
13. Teachers Knowledgeable	0	1	4	65	30
14. Ensure Quality	0	2	7	47	44
15. Effective Approaches	0	1	13	52	34
16. Strategies	0	1	13	44	42
17. Focus on Goals & Curriculum	0	4	19	50	27
18. Performance Assessments	1	6	18	54	21
19. Staff Development	6	23	33	29	9

Table 59. Teachers' Perceptions of Their Personal Professional Development N=150

Question #20 ***Over the past year, I have…***	Percentage of responses
Engaged in a mentoring relationship with another teacher.	49%
Participated in joint planning activities with other SAGE teachers.	96%
Collaborated with other teachers in delivering lessons.	88%
Collaborated with other teachers in evaluating student progress.	90%
Participated in a study group or on-line network.	24%
Collaborated in school-wide instructional initiatives or themes.	68%
Collaborated with other schools or institutions.	29%
Conducted research connected to my teaching.	26%
Attended a professional conference or skill-building workshop.	85%
Attended a workshop, seminar or retreat focused on diversity or human relations training.	28%
Attended a workshop, seminar or retreat focused on teaching smaller classes.	15%
Taken a course for graduate of CEU credit.	46%
Question 21 ***Do you have a personal formal, written professional development plan?***	
Yes	46%
No	54%
Question 22 ***Which of the following statements most accurately reflects the content of your professional development plan?***	
It was determined primarily by me	36%
It was determined in consultation with school administrators.	14%
It was determined in consultations with district administrators.	1%
It was determined primarily by school and /or district administrators	6%

In terms of the content of staff development programs in SAGE schools, most teachers agree that staff development in their schools is on-going, it enjoys financial as well as professional and community support, and it brings about change on the part of teachers. There is limited agreement, however, about the extent to which it occurs in joint learning activities inside or outside of the school.

The findings regrding the process of staff development are that teachers strongly agree that they are knowledgeable about child learning and development, are able to provide quality education for all students, and have high expectations for all students. The areas in which there is the least agreement are that teachers are involved in decisions concerning staff development,

staff development includes practical activities, and staff development focuses on teaching in reduced size classrooms.

Teachers' perceptions of their own development are contained in Table 59. These results show that almost half of the responding teachers have professional development plans, and of these about 60 percent are determined by the teachers themselves. The teachers' professional development over the past year consisted, primarily, of joint planning with other SAGE teachers, collaborating with other teachers in delivering lessons and evaluating student progress, and attending professional conferences or skills workshops. Those professional development activities mentioned the least were attending workshops on reduced size classrooms or on diversity, participating in study sessions or collaborating with other schools, and conducting research regarding their teaching.

Family Involvement and Lighted Schoolhouse

The extent to which SAGE school parents are involved in the education of their children is reported in Table 60 for 1996-97 and 1997-98. It can be seen that parent involvement is relatively stable over the two-year period. Parent-school contacts occur most frequently through teacher notes, teacher and parent conversations before or after school, and telephone calls. Home visits and weekly progress reports occur infrequently.

Table 60. Questionnaire Results for Family Involvement

ITEM	1996-97 (N=212)	1997-98 (N=315)
Class Newsletter	71	62
Weekly progress report—requiring parent signature	24	28
Weekly progress report—not requiring parent signature	11	12
Notes sent home	98	93
Conversations with parents	95	94
Parental visits to school	74	76
Telephone calls	92	89
Home visits	10	14
Structured after school activities	(not used)	33

Table 61 shows lighted schoolhouse data from the first two years of SAGE and the year prior to SAGE. Data regarding lighted schoolhouse activities existing prior to SAGE were obtained from the Baseline Data Questionnaire administered in May 1996, and from the school contracts completed for DPI prior to a school's enrollment in the SAGE program. Data for the first year of SAGE, 1996-97, were obtained from principal interviews in addition to the year-end reports required by DPI. Data for 1997-98 were obtained from the year-end reports required by DPI.

Schools report a progressive increase in the number of lighthouse activities and a corresponding increase in the number of participants. Principal interview data from 1996-97 suggested that SAGE schools took responsibility for the conception and operation of the lighted schoolhouse activities (as opposed to activities initiated by parents or community volunteers). Those interviews also suggested that schools did not focus heavily on their lighted schoolhouse activities in the first year of SAGE implementation. However, the data show that in nearly every area, the number of lighthouse activities and the number of participants have increased steadily since SAGE implementation.

Table 61. SAGE Schools' Lighted Schoolhouse Activities for Students and for Families

	Pre-SAGE	1996-97	1997-98	Change from Pre-SAGE to 1996-97	Change from 1996-97 to 1997-98	Total Change from Pre-SAGE to 1998
Student Activities						
Child Care	12	13	17	4	1	5
Health Clinic	8	11	13	1	4	5
Breakfast	15	18	26	8	3	11
Tutoring	16	17	20	3	1	4
Homework Helpline	4	2	10	8	-2	6
Extended Library Hours	6	8	12	4	2	6
Community Recreation	14	19	25	6	5	11
Girl Scouts/Brownies	15	12	22	10	-3	7
Boy Scouts/Cub Scouts	12	12	23	10	1	11
Music Lessons	5	9	9	0	4	4
Summer Reading Program	0	0	13	13	0	13
Head Start	6	5	9	4	-1	3
Family Activities						
Social Services	1	4	5	1	3	4
Health Clinic	4	8	7	-1	4	3
Family Resource Center	4	5	7	2	1	3
Adult Tech. Ed.	5	4	8	4	-1	3
GED Preparation	3	3	5	2	0	2
Extended Library Hours	5	5	11	5	1	6
Community Recreation	9	11	20	1	5	6
FAST	5	5	3	-2	0	-2
PTA/PTO	20	17	25	8	-3	5
Community Education	2	8	13	5	6	11
Even Start Literacy	1	2	5	3	1	4
Parent/Community Advisory	15	18	21	3	3	6
Other	7	36	31	4	20	24
Totals						
Total Programs	202	213	329			
Total Participants	11,766	15,678	20,796			

DISCUSSION: MAJOR FINDINGS, LIMITATIONS, AND FUTURE REPORTS

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, provide before- and after-school activities for students and community members, and 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. This is the second of five annual evaluation reports.

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 added third-grade classrooms at those schools. The SAGE evaluation will continue to assess the impact of the program in these schools through the 2000-01 school year.

To measure academic achievement first-grade students in SAGE schools and in a group of comparison schools were tested in October 1997 and again in May 1998 using the Comprehensive Test of Basic Skills (CTBS) Complete Battery, Terra Nova edition, Level 10 (fall) and Level 11 (spring). Second-grade students were administered Level 12 in May 1998. It was decided that a standardized test at the kindergarten level was not an appropriate evaluation measure. Therefore, standardized tests are not administered to kindergarten students as part of the SAGE evaluation.

Major Findings

The Achievement Effect of Class Size Reduction

First grade

- Students in SAGE classrooms achieved significantly higher scores than students in comparison school classrooms in all tested areas: mathematics, reading, and language arts. The total scores of students in SAGE classrooms were also significantly higher than those of comparison school students (see Tables 14-18, pages 24 and 25).
- African-American SAGE students, the programs largest minority group, scored lower than African-American students in comparison schools in the fall pre-test, but made significantly larger gains than comparison school students from pre-test to post-test. The test scores of African-American students in SAGE classrooms surpassed those of African-American students in comparison school classrooms on the spring post-test.
- African-American students in SAGE classrooms achieved greater gains on the CTBS total score than white SAGE students from pre-test to post-test, reducing the achievement gap. In contrast, African-American students in comparison school classrooms achieved lessor gains and the gap in achievement between African-American students and white students widened.

Second Grade

- The achievement advantage of students in SAGE first-grade classrooms in 1996-97 appears to be maintained in second grade in 1997-98. The advantage, however, does not appear to have increased significantly.

Achievement Effects Related to SAGE Classroom Organization

Classrooms in the SAGE program achieve a 15:1 student-teacher ratio in several ways. Most SAGE classrooms have roughly one teacher and fifteen students. Other classroom types are also utilized, e.g., two teachers sharing a single classroom with 30 students but each teacher

teaching 15 students; two teachers team teaching roughly 30 students; and classrooms with roughly 30 students that utilize a "floating teacher" to help teach selected subjects. There are not enough classrooms of any single type of these latter types of classroom organization to do separate analyses reliably. However, it is possible to compare the performance of classrooms with a 15:1 student-teacher ratio to classrooms with a 30:2 student teacher ratio. The results of analyses of first-grade classrooms found no achievement advantage for the one teacher to fifteen students form of classroom organization on any of the CTBS sub-tests.

Reduced Class Size and Life in SAGE Classrooms

During the 1996–97 and 1997-98 school years, members of the SAGE evaluation team conducted teacher interviews, made classroom observations, analyzed logs kept by SAGE teachers describing their classroom activities, and tabulated the results of teacher questionnaires on a variety of teaching and learning topics. Taken together, these data provide a picture of life in SAGE classrooms which includes these features:

- Teachers have greater knowledge of each of their students.
- Little time is required to manage the class resulting in more time for instruction.
- Individualized instruction is pervasive and constant.
- The type of instruction used is mainly teacher-centered (e.g., students listen, practice, answer, etc). Teachers also used hands on activities with some regularity, but student-centered teaching (e.g., problem solving, cooperative groups, etc.) is used infrequently.

Other SAGE Interventions

In addition to reducing class size the SAGE program calls for participating schools to develop rigorous academic curricula, implement programs of staff development and professional accountability, and be open extended hours for lighted schoolhouse activities.

Rigorous Curriculum

Based on teachers' perceptions, national standards in mathematics, reading, and language arts are being implemented in SAGE classrooms. The congruence between the standard and teaching is greater in reading-language arts than in mathematics, however, specific areas in which agreement is comparatively low are areas often thought to be more appropriate for older children.

Professional Development

Staff development appears to be an important feature of SAGE schools. Most teachers view it as an ongoing, regular program that produces changes in classroom practices. Although many report that they are not involved in making decisions about school staff development, are not assessed in terms of the use of innovations, and are not given strategies to use in reduced size classes, they see themselves working collaboratively with their colleagues in making decisions about lessons and evaluating students.

Lighted Schoolhouse

In 1997-98 SAGE schools reported a rise in lighted schoolhouse activities when compared to 1996-97 and to the year prior to their entry to the SAGE program. Over the two-year period, SAGE schools reported a 78% increase in the number of activities offered and a 77% increase in the number of people participating in these activities.

Limitations

When considering the results of the 1997-98 SAGE evaluation several factors should be kept in mind:

- The number of schools in the comparison group pool was reduced from 16 in 1996-97 to 14 in 1997-98. One school converted from a comparison school to a SAGE school. Two

additional schools withdrew for other reasons. This problem has been addressed for the 1998-99 academic year by the addition of three comparison schools.

- Second grade results may have been influenced by three factors. During 1996-97 a considerable number of SAGE first graders achieved perfect scores on the spring post-test. This had the potential effect of placing a "ceiling" on the gains reported for SAGE first-grade students. Conversely, what was a "ceiling" in 1996-97 became a "floor" for the scores of this group of SAGE students in second grade. It is not possible to know to what extent this phenomenon had an impact on the 1997-98 SAGE second-grade achievement results. A second factor that may have influenced the second-grade results reported for SAGE students was that, because of uncertainties over funding for the second grade, nine of the thirty SAGE schools did not implement the program in second grade until after the start of the school year. In some cases implementation was delayed until January 1998. (See Table 62 for a list of schools implementing SAGE in second grade after the start of the 1997-98 school year.) Finally, the impact on class achievement scores of non-SAGE students entering the SAGE program for the first time in second grade is unknown.

Table 62. Second Grade Implementation of SAGE

SCHOOL	IMPLEMENTED SECOND GRADE
Prentice – Ogema	10/31/97
Tripoli	10/31/97
Glidden – Glidden	11/20/97
Milwaukee-Maryland	9/30/97
Sherman	12/1/97
Carlton	Class 1 – 11/03/97
	Class 2 – 11/10/97
Maple Tree	Class 1 – 11/15/97
	Class 2 – 01/05/98
	Class 3 – 01/05/98
Longfellow	Class 1 – 11/14/97
	Class 2 – 11/24/97 (long term sub)
	Class 3 – 12/09/97
Wisconsin Conservatory	January 1998

- The longitudinal results reported are based on regression equations using individual level data. Since SAGE students do not move from grade to grade as part of an intact cohort it was not possible to use HLM analyses to compare the performance of SAGE second-grade classrooms to that of SAGE first-grade classrooms the preceding year.
- The analyses of the impact of different types of SAGE classroom organization on achievement outcomes must be considered preliminary. The analyses were conducted using only first-grade classroom organization data. Analyses conducted in subsequent years will use classroom organization data for first-, second-, and third-grade SAGE classrooms. These analyses are necessary to confirm this year's findings.
- The qualitative data reported here do not draw distinctions between teacher behavior in kindergarten, and first-, and second-grade classrooms. Additional analyses will be necessary to ascertain whether grade level differences in teacher behavior can be identified.

Future SAGE Evaluation Reports

- Smaller classes in the SAGE program have a significant effect on student achievement in the first grade. The data from the 1996-97 evaluation and the 1997-98 evaluation indicate that this finding is robust. In future reports the achievement effects of the SAGE program beyond first grade should emerge much more clearly than in this report. The problems created by using the CTBS Complete Battery, Terra Nova edition, Level 10 as a first-grade post-test during the 1996-97 school year will continue to make interpretation of the achievement results of that cohort of SAGE first-grade students problematic as they move through second and third grade. Much clearer results should emerge as the SAGE 1997-98 and 1998-99 first-grade cohorts move through the grades.
- The 1998-99 report will consider the impact of attendance in SAGE kindergarten on

subsequent SAGE student achievement.

- Although the SAGE program specifies that participating schools reduce class size to a 15:1 student-teacher classroom ratio, the average class size in SAGE schools in 1997-1998 varied. Meta-analyses of the effects of reduced class size by Glass and Smith (1978) suggest that even a slight increase in class size has an impact on achievement. Future analyses of SAGE data will consider the effects of this increased class size.
- The impact on class achievement results of non-SAGE students entering second and third grade for the first time will be considered.
- Several aspects of life in SAGE schools and classrooms will emerge more vividly in subsequent reports. Examples of areas being investigated include how teacher behaviors do or do not change over time in the SAGE program; how teachers in different grades respond to smaller classes similarly and differently; what impact lighted schoolhouse programs have on SAGE schools and on school community relations; how staff development and professional accountability programs evolve in SAGE schools; and the impact of non-SAGE students entering the program in second and third grade.
- Future reports will continue to study whether or not different forms of SAGE classroom organization have a differential impact on student achievement. If the 1997-98 findings are sustained the implications are considerable for school districts such as Milwaukee that wish to reduce class size but do not have the funds to construct additional school space.

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.

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

CTB/McGraw-Hill (1991). *Comprehensive Test of Basic Skills 4th Edition*. CTB Macmillan/McGraw-Hill.

Glass, G. and 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.

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.

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

Interview Guide
SAGE
Teacher Interview

1. Describe the extent to which your teaching is changing as a result of having fewer students.
Probe for in-depth *descriptions, details, examples,* etc., of the following:

Teaching time

a) More time on instruction?
b) Less discipline required?

Teaching

c) More individualization of instruction?
d) More use of student-centered activities (cooperative groups, interest centers, manipulatives, etc.)?
e) Greater number of learning activities used?
f) More opportunity for students choice of activites?
g) More opportunity for students to problem solve, create, etc?

Content

h) More content (subject matter and skills) covered?
i) Content covered in greater depth?

Students

j) Greater knowledge of students' abilities, needs, personalities, etc?
k) Students more attentive and engaged?
l) Studnets participate in more discussions, volunteer comments more, etc?

Teacher

m) Greater enthusiasm for and enjoyment of teaching?
n) Greater effort devoted to planning and teaching?

2. Describe the extent to which you believe your students' learning has improved as a result of being in a class with fewer students.
Probe for in-depth *descriptions, details, examples, etc.,* of the following:

a) Improved reading/language arts?
b) Improved mathematics?

3. Describe changes you anticipate in your teaching during the second year of SAGE.

Appendix B

Classroom Observation Guide

1. Purpose of observations
   To develop a descriptive, chronological account of selcect classrooms.

2. Classroom observation guide

   Substantive Notes

   a) Setting: Where is it taking place?
      Space: desks, furniture, objects, displays, interest centers, etc.
      Atmosphere: suportive, hostile, etc.

   b) People: Who is involved?
      Appearance
      Roles
      Grouping: total class, small groups, individuals

   c) Events: What's going on?
      Learning activities
      Participants' behavior
      Participants' talk (verbatim, if significant)
      Objectives and specific subject matter or/and skill
      Materials and resources

   Reflective Notes

   a) Analysis: speculate on themes, patterns, etc.
   b) Method: observation problems, next steps, etc.
   c) Clarifications: points that need to be clarified and explained
   d) Frame of mind: observer's biases, etc.

SAGE Classroom Observation Categories

Individualization

I1.	Monitoring	T moves about room to check on progress of students' work.
I2.	Grouping	T divides class into subgroups or pulls out a student for special attention.
I3.	Choice	T permits students to create own learning activities, select learning centers, etc.
I4.	Help	T offers feedback critique, assistance, guidance, scaffolding, help to student.
I5.	Whole Class	T provides whole class instruction.
All.	All Children	T enables All children in the class to participate in a specific activity or discussion.

Engagement

E1.	Listening	S listen to teacher directions, demonstrations, lectures, explanations, stories, etc.
E2.	Practicing	S work at their seats to complete workbook exercises, board work, worksheets, read textbooks, read trade books, use flash cards, etc.
E3.	Responding	S respond orally to teacher questions, follow teacher direction to write on the chalk board, point to an object, read aloud, recite in unison, etc. (Identify each participant, e.g., S1, S2, S3.)
E4.	Gaming	S play educational or recreational games, role play, dramatize, sing.
E5.	Manipulating	S manipulate blocks, markers, objects, etc.
E6.	Creating	S draw, paint, make displays, work on projects, write stories, etc.
E7.	Dialoguing	S engage in discussion with other students and/or the teacher in which positions are stated, questioned, critiqued, clarified, etc.
E8.	Problem Solving	S engage in investigation, inquiry, experimentation by formulating questions, drawing conclusions, collecting data, etc.
E9.	Reporting	S share, present, report on accomplishments, ideas, etc.
E10.	Reflecting	S evaluate their knowledge and skill based on teacher critique, experiential feedback, etc.
E11.	Initiating.	S volunteers own ideas, perception, understanding, questions to class interaction.

Management

M1.	Praise	T gives oral praise, stickers, prizes, etc., for academic achievement or appropriate behavior.
M2.	Reproof	T gives oral reproof, isolates a student, issues a threat for inappropriate behavior.
M3.	Remind	T reminds students of class rules, procedures, etc., regarding appropriate behavior.
M4.	Warms	T personalizes learning by relating topics, ideas, etc, to students' lives, telling jokes, sharing own experiences, laughing, etc.
M5.	Cools	T turns students off to learning by ignoring students, making cutting comments, sarcasm, etc.
M8.	Peer	T allows students to develop socialization skills in areas relating to problem solving amongst peers.
M11.	Permits	T permits students to make choices regarding behavior (bathroom, water, other physical behaviors.)

Appendix C

Teacher Log

Time Period	Type	Grouping	Content	Student Learning Activities (3)
7:30-7:45				
7:45-8:00				
8:00-8:15				
8:15-8:30				
8:30-8:45				
8:45-9:00				
9:00-9:15				
9:15-9:30				
9:30-9:45				
9:45-10:00				
10:00-10:15				
10:15-10:30				
10:30-10:45				
10:45-11:00				
11:00-11:15				
11:15-11:30				
11:30-11:45				
11:45-12:00				
12:00-12:15				
12:15-12:30				
12:30-12:45				
12:45-1:00				
1:00-1:15				
1:15-1:30				
1:30-1:45				
1:45-2:00				
2:00-2:15				
2:15-2:30				
2:30-2:45				
2:45-3:00				
3:00-3:15				
3:15-3:30				
3:30-3:45				
3:45-4:00				
4:00-4:15				
4:15-4:30				

Type
0. Routines (attendance, lunch money, etc.)
1. Instructional time (teaching and learning)
2. Planning and Evaluation(lesson plans, grading, etc.)
3. Personal time (snacks, phone, etc.)
4. Housekeeping/Clerical (xeroxing, supplies, etc.)

Grouping
5. Whole Group
6. Small Group
7. Individual
8. Combined Classes

Content
9. Reading/Language Arts
10. Mathematics
11. Integrated (thematic, authentic)
12. Other (science, music, etc.)

Student Learning Activities
13. Listening (information, modeling etc. from teacher)
14. Practicing (practice skills in seat, board work, etc.)
15. Creating (making, dramatizing, etc.)
16. Receiving help (assistance from teacher)
17. Using manipulatives (games, objects, etc.)
18. Dialoguing (discussing, reacting)
19. Problem Solving (collecting data, etc.)
20. Receiving critique (feedback, guidance, coaching)
21. Reflecting (self-evaluation)
22. Responding (responding to questions, directions, etc.)
23. Reporting (sharing accomplishments)
24. Recitation (reading aloud, reciting in unison)