Alex Molnar, Philip Smith, John Zahorik, Amanda Palmer, Anke Halbach, Karen Ehrle

Evaluating the SAGE Program: A Pilot Program in Targeted Pupil-Teacher Reduction in Wisconsin

Educational Evaluation and Policy Analysis, Vol. 21 No. 2
Summer 1999 Pp. 165-177

Educational Evaluation and Policy Analysis
Summer 1999, Vol. 21, No. 2, pp. 165–177

Evaluating the SAGE Program: A Pilot Program in Targeted Pupil-Teacher Reduction in Wisconsin

Alex Molnar Philip Smith John Zahorik Amanda Palmer
Anke Halbach Karen Ehrle
University of Wisconsin–Milwaukee

Wisconsin's Student Achievement Guarantee in Education (SAGE) program was designed as a 5-year K–3 pilot project that began in the 1996–97 school year. The program requires that participating schools implement 4 interventions including reducing the pupil-teacher ratio within classrooms to 15 students per teacher. The SAGE evaluation uses a quasi-experimental, comparative change design utilizing descriptive statistics, linear regression, and hierarchical linear modeling. In addition, qualitative analyses of life in SAGE schools and classrooms are conducted. Results of the 1996–97 and 1997–98 first grade data reveal findings consistent with the Tennessee STAR class size experiment. Also, individualization emerged as a key characteristic of instruction in SAGE classrooms.

Wisconsin's Student Achievement Guarantee in Education (SAGE) program was designed as a 5-year pilot project that began in the 1996–97 school year. The program requires that participating schools implement the following four interventions: (a) reduce the pupil-teacher ratio within a classroom to 15 students per teacher beginning with kindergarten and first grade in the 1996–97 school year (second grade was added in 1997–98 and third grade in 1998–99), (b) establish "lighted schoolhouses" open from early in the morning until late in the evening, (c) develop "rigorous" curricula, and (d) create a system of staff development and professional accountability.

The SAGE program was targeted toward schools with a high proportion of students living in poverty. School districts in Wisconsin that had at least one school with 50% of children or more living below the poverty level were eligible to apply for participation in SAGE. Within those districts, any school with 30% of students or more below the poverty level was eligible to become a SAGE school. Eligible districts were allowed only one SAGE school with the exception of the Milwaukee school district, which was allowed up to 10 SAGE schools. No schools were to be allowed to enter the program after the initial eligibility period. Funding was set at a maximum of $2,000 per low-income student enrolled in SAGE classrooms (K–3). No school district that applied for SAGE was rejected, and during the 1996–97 school year 30 schools in 21 school districts, including 7 in Milwaukee, began the program in K–1. Second grade was added for these schools in 1997–98 and third grade in 1998–99.

The SAGE legislation mandated an annual evaluation of the effects of the program. We report the results for the evaluation of the first 2 years of the SAGE program for first grade. The SAGE evaluation is based on comparisons of achievement in the 30 schools that entered the program in the autumn of 1996 and a group of 14–17 preselected comparison schools with similar student and school characteristics. Achievement tests were administered in the SAGE and comparison schools at the beginning and end of the first grade. We also report on the results of teacher surveys and interviews of teachers in the smaller SAGE classrooms.

Of the four SAGE interventions, only reduced pupil-teacher ratio within a classroom was uniformly and immediately implemented across all program schools. This reduction was accomplished

mainly by assigning 15 or fewer students to a teacher within a single classroom. However, a number of alternate configurations were also adopted. They included classes of approximately 30 with two teachers in a single classroom and other configurations. Pupil-teacher ratios averaged between 12:1 and 15:1 in SAGE classrooms and between 21:1 and 25:1 in comparison schools. These differences were larger than class size reductions in the Tennessee experiment and would be predicted to have measurable effects based on the Tennessee evidence.

In contrast, the other SAGE interventions would probably not be expected to show significant effects for several reasons. There were no measurable guidelines for determining "a rigorous curriculum," a system of "professional development and accountability," or specific "lighted schoolhouse" programs/procedures to implement. These provisions of the program were implemented by schools with considerable variation and sometimes with considerable delay. In addition, meaningful changes in curriculum and professional development usually take years, so any impact from these provisions would be neither uniform nor immediate. The lighted schoolhouse was also not uniformly and immediately implemented in all schools. Moreover, no specific instructional initiatives were implemented in the before school and after school hours. So while we cannot analytically rule out that a part of the measured effects might be from these provisions, we expect that reductions in pupil-teacher ratio within classrooms would be the primary cause of effects from the SAGE initiative. However, the effects measured here and attributed to class size reductions should be treated as upper limits until more analysis can be done.

Evaluation Design

To determine the impact of SAGE pupil-teacher reductions on student achievement, the SAGE evaluation uses a quasi-experimental, comparative change design. The quasi-experimental design was used because it was not possible to randomly assign students and teachers to classrooms, to keep classroom cohorts intact from year to year, and to control the class size requirement in other ways. Fiscal constraints and the lack of incentives for participation by comparison schools prevented the use of matched-pair comparison schools. Therefore, SAGE classrooms are compared with a set of comparison schools that have normal class sizes,

TABLE 1
Characteristics of SAGE and Comparison Students 1996–97 and 1997–98, All Grades

	Percentage of students 1996–97		Percentage of students 1997–98	
Characteristic	SAGE	Comparison	SAGE	Comparison
Gender				
Female	48.6	49.4	48.4	48.5
Male	51.4	50.6	51.6	51.5
Race/ethnicity				
African American	24.8	32.9	26.3	24.7
Asian	5.7	5.5	5.2	5.6
Hispanic	6.6	8.0	6.5	10.0
Native American	11.7	1.4	10.3	1.5
White	48.8	49.0	43.8	52.2
Other	1.6	2.7	2.0	2.3
Subsidized lunch eligibility				
Free	57.7	49.4	54.0	43.4
Reduced	10.9	9.9	10.6	8.9
Not eligible	31.4	40.7	35.4	47.7
Repeating grade	3.2	2.6	2.7	2.0
English as second language	8.2	4.9	7.9	6.4
Referred to M-Team	13.6	9.2	9.6	6.8
Exceptional education need	13.1	9.7	10.0	7.1
Individualized education plan	8.2	5.5	8.8	5.6

TABLE 2
Enrollment Changes in SAGE and Comparison Schools by School Year for All Grades

	SAGE				Comparison			
	1996–97		1997–98		1996–97		1997–98	
	Number	%	Number	%	Number	%	Number	%
Ongoing	2,943	81.4	2,455	42.3	1,706	85.3	1,402	44.3
Withdrew	397	11.0	1,093	18.8	178	8.9	589	18.6
Enrolled	274	7.6	2,262	39.0	115	5.8	1,175	37.0

are from districts participating in the SAGE program, and, as a group, resemble SAGE schools in family income, achievement in reading, K–3 enrollment, and racial composition. In 1996–97 there were 17 comparison schools, and in 1997–98 there were 14. Gender, race, and other characteristics of students in SAGE and comparison schools are displayed in Table 1. The students in SAGE and comparison schools had similar overall minority percentages in 1996–97, but SAGE schools had smaller proportions of Black and Hispanic students and greater proportions of Native American students. In 1997–98, SAGE schools had somewhat larger minority enrollment than comparison schools, mainly as a result of differences in Native American population. SAGE schools also had a somewhat greater proportion of students with free or reduced-lunch eligibility and students with exceptional need or individualized education plans.

Many students withdraw from SAGE and comparison schools during the year, while others enroll. As Table 2 shows, student turnover, as measured by those leaving and those enrolling during the year, is somewhat greater in SAGE schools. Table 3 shows that SAGE schools are significantly larger than comparison schools.

These differences in SAGE and comparison schools do not result in significant discrepancies in the test scores of beginning first-grade students. Table 4 shows the results of these achievement tests. Inferential tests comparing pretest (fall) scores for the SAGE and comparison groups showed no significant differences for either the 1996–97 or 1997–98 first-grade students.

Classroom Interventions

SAGE schools reduced pupil-teacher ratio 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 occurred 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 *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 *two-teacher team* classroom is a class where two teachers work collaboratively to teach as many as 30 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 used in the SAGE program but to a very 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

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

	SAGE		Comparison	
	1996–97	1997–98	1996–97	1997–98
Kindergarten	1,494	1,524	820	676
First grade	1,723	1,567	1,001	985
Total	3,217	3,091	1,821	1,661

TABLE 4
Means (and Standard Deviations) for First-Grade Achievement Scores

	Fall 1996 Pretest		Fall 1997 Pretest	
	SAGE	COMP	SAGE	COMP
Reading	531.04	533.66	533.35	535.06
	(36.88)	(38.24)	(36.43)	(36.18)
Language arts	527.29	529.15	530.50	528.97
	(43.59)	(44.27)	(43.78)	(43.39)
Math	490.12	489.76	492.34	493.02
	(40.21)	(40.56)	(42.51)	(38.38)
Total	516.49	518.03	519.06	519.51
	(35.08)	(35.43)	(35.34)	(33.35)

Note. COMP = comparison.

afternoon, and the *three-teacher team* classroom, where there are 45 students taught collaboratively by three teachers.

The types of classrooms are displayed in Table 5. SAGE classes ranged in number of students from 7 to 38. A few SAGE classrooms exceeded the 15:1 student-teacher ratio, but only by 1 or 2 students. The percentage of students per one teacher and average number of students per teacher are shown in Table 6.

Data Collection Instruments

To provide information about the processes and products of the SAGE program for 1996–97 and 1997–98, a number of instruments were used as part of the evaluation. A description of the test and nontest instruments used in the evaluation presented here follows. We also collected data from principals, derived information from teacher assessments of student participation, and did classroom observations. Analyses of these data are presented in two annual evaluation reports and will be the basis of future investigations (Maier, Molnar, Smith, & Zahorik, 1997; Molnar, Smith, & Zahorik, 1998).

1. *Comprehensive Test of Basic Skills.* 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 was administered 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 reading, language arts, and mathematics subtests. The CTBS was chosen as an achievement measure because it is derived from an item response theory 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. All inferential results for the Terra Nova are reported in standard score format.

2. *Student profiles.* These instruments were completed by schools in October and May and provided demographic and other data on each SAGE school and comparison school student. The variables derived from this instrument included race/ethnicity, socioeconomic status (SES), and SAGE participation. Race/ethnicity was dummy coded 1 if a student was of a certain race/ethnicity and 0 if not. Variables were included for African American and White, along with a residual race category ("other"). Eligibility for subsidized lunch was used as an indicator of SES. This variable was coded 0 if a student was ineligible, 1 if a student was eligible for reduced-price lunch, and 2 if a student was eligible for free lunch (this variable was assumed to be interval level). Finally, a dummy variable was used for SAGE participation. This variable was coded 0 if a student was from a comparison school and 1 if a student was from a SAGE school.

3. *Classroom organization profile.* Completed in October, this instrument was used to record how

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

	Regular		Shared space		2-teacher team		Floating teacher		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	2	4	24	22	3	2	0	0	1	0
First grade	61	84	8	8	18	23	7	2	2	0	0	1

TABLE 6
Student-Teacher Ratio for SAGE and Comparison Classrooms

	SAGE classrooms (%)		Comparison classrooms (%)	
Number of students per teacher	1996–97 first grade	1997–98 first grade	1996–97 first grade	1997–98 first grade
7–13 students	18	28	9	12
14–16 students	62	64	0	7
17+ students	20	8	91	81
Average class size	14.63	13.47	24.53	22.42

SAGE schools attained a 15:1 student-teacher ratio.

4. *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.

5. *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.

Analysis of Student Achievement Outcomes

Statistical Analysis

The SAGE evaluation design uses descriptive statistics, linear regression, and hierarchical linear modeling. Descriptive statistics are presented for all students and separately for African-American and White students. Regression models were used to control for differences in individual student characteristics while measuring SAGE effects. Hierarchical linear modeling was used to control for class-level characteristics as well as individual-level characteristics; it also involved an actual classroom pupil-teacher ratio rather than a dummy variable for SAGE.

First-Grade Results: Descriptive Statistics

The number of first graders for whom valid test scores were available was less than the total number of first-grade students. There are four main explanations for this. First, the evaluation team presented schools with the option of allowing Exceptional Educational Needs (EEN) and English-as-a-second-language students to take the test, even though the test may be inappropriate for these students. These scores were invalidated on the basis of 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 pretest scores but no posttest 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 were unavailable for these students. Finally, some of the students did not complete the pretest, posttest, or both the pretest and posttest. The numbers of valid test scores for the first-grade students in both cohorts are presented in Table 7.

Table 8 provides descriptive statistics for the CTBS from the pretest and posttest results for both of the first-grade cohorts. Inferential tests comparing pretest (fall) scores for the SAGE and comparison groups showed no significant differences for either cohort. Tests comparing posttest (spring) scores for the SAGE and comparison groups revealed significant differences ($p < .05$) on all test scores for both cohorts.

First-Grade Results: Regression Analysis

The effect of the SAGE program on student achievement, controlling for other factors, was tested through a series of regression models for each subtest and for total scale scores. A summary of the results of the regression analyses is presented in Table 9. Pretest score was a consistent and statistically significant predictor of performance on all subtests and total scale score. Days absent and lunch eligibility emerged as consistent and statistically significant predictors on nearly all tests. For all tests, attending a SAGE school emerged as a significant predictor of student achievement on the

TABLE 7
Number and Percentage of First-Grade Students With Valid Test Scores

	1996–97				1997–98			
	Number		Percentage		Number		Percentage	
	SAGE	COMP	SAGE	COMP	SAGE	COMP	SAGE	COMP
Reading	1,575	941	91	94	1,318	844	84	86
Language arts	1,575	941	91	94	1,319	844	84	86
Mathematics	1,566	928	91	93	1,334	841	85	85
Total	1,552	918	90	92	1,310	829	84	84

Note. COMP = comparison.

posttest. The magnitude of the effect of SAGE on student achievement ranged from 3 to 7 points, depending on the CTBS subtest. The overall effects of 3.3 points in 1996–97 and 6.3 points in 1997–98 translate to approximately 0.1 and 0.2 standard deviation score gains. Because of the test level used in the 1996–97 first-grade posttest, there was a noticeable ceiling effect on these results, especially for the SAGE participants. The smaller SAGE effects for the 1996–97 cohort are probably due to this test ceiling effect.

First-Grade Results: White and African-American Students

The Tennessee STAR experiment found that African-American students benefited from reduced class sizes even more than White students. For this reason, the performance of African-American students in the SAGE program was of particular interest. Among minority students in SAGE and comparison schools, African Americans clearly represented the largest group with valid test scores: roughly 25% of all SAGE students and 28% of all comparison school students. Table 10 provides mean CTBS pretest and posttest scores for African-American students in SAGE and comparison schools. African-American SAGE students scored lower on the pretest than African-American comparison students on every subtest except for reading in 1997–98. However, on the posttest, African-American SAGE students scored significantly higher than African-American comparison school students on every subtest and had significantly higher total scale scores.

African-American students had significantly lower CTBS pretest total scale scores than White students, as shown in Table 11. This result held for both SAGE and comparison schools, although the gap between African-American and White students was larger in SAGE schools. African-American SAGE students achieved greater gains in terms of CTBS total scale score than White SAGE students from pretest to posttest, closing the achievement gap (although the gap remained statistically significant). In contrast, African Americans in comparison schools achieved lesser gains, and in these

TABLE 8
Means (and Standard Deviations) for First-Grade Cohorts

	Fall 1996 pretest		Spring 1997 posttest		Fall 1997 pretest		Spring 1998 posttest	
	SAGE	COMP	SAGE	COMP	SAGE	COMP	SAGE	COMP
Reading	531.04	533.66	582.33	578.66	533.35	535.06	580.33	570.80
	(36.88)	(38.24)	(37.17)	(39.99)	(36.43)	(36.18)	(41.33)	(45.52)
Language arts	527.29	529.15	581.09	575.31	530.50	528.97	586.02	573.98
	(43.59)	(44.27)	(39.45)	(42.71)	(43.78)	(43.39)	(45.33)	(46.84)
Math	490.12	489.76	545.56	538.27	492.34	493.02	538.63	525.14
	(40.21)	(40.56)	(42.64)	(44.50)	(42.51)	(38.38)	(40.09)	(42.53)
Total	516.49	518.03	569.90	564.50	519.06	519.51	568.63	556.87
	(35.08)	(35.43)	(33.93)	(35.58)	(35.34)	(33.35)	(36.66)	(38.83)

Note. COMP = comparison.

TABLE 9
Scale Score Regression: First-Grade Block Three Beta Coefficients

	1996–1997				1997–1998			
	Reading	Language arts	Math	Total	Reading	Language arts	Math	Total
Pretest score	0.34*	0.41*	0.62*	0.61*	0.54*	0.55*	0.61*	0.76*
Days absent	–0.38*	–0.42*	–0.37*	–0.35*	–0.16	–0.33*	–2.15*	–2.56*
Lunch eligibility	–4.24*	–3.42*	–1.32	–2.15*	–2.15*	–2.06	–2.56*	–0.62
African American	–8.82	–6.17	–0.37*	–8.19*	2.80	2.50	1.44	3.72*
White	–0.61	5.75	0.07	–0.37	7.37*	5.80*	3.24	3.20*
SAGE	2.74*	3.85*	3.88*	3.30*	6.98*	7.25*	7.06*	6.33*
Constant	378.14	365.51	245.95	256.06	296.11	304.36	247.05	176.57
Adjusted R^2	0.24	0.29	0.41	0.48	0.27	0.34	0.44	0.55
SE	32.84	33.60	33.26	24.54	35.50	36.30	29.88	24.34

*$p < .05$.

schools the achievement gap widened.

First-Grade Results: Hierarchical Linear Modeling

The SAGE data include student- and classroom-level information, 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). Hierarchical linear modeling (HLM) (Bryk & Raudenbush, 1992) was specifically designed to accommodate these types of data structures. HLM was used with the SAGE data to determine whether different results emerge when both student and classroom effects are included. In these models, variables associated with individual students are referred to as Level 1 variables, and those associated with classrooms are referred to as Level 2 variables.

For all analyses, the Level 1 variables were pretest achievement scores and SES measured as eligibility for subsidized lunch. The 1996–97 analysis also used "attendance" as a Level 1 variable; however, this variable was not used in 1997–98 because these data were not reported by several districts. Three different Level 2 models are reported here: The first includes actual classroom pupil-teacher ratio as the only Level 2 variable, the second adds the SAGE dummy variable, and the third includes class SES and actual classroom pupil-teacher ratio.

Table 12 provides a summary of the effects. The

TABLE 10
African-American Students' First-Grade Scores, by SAGE or Comparison School

	1996–1997		1997–1998	
Scale score	SAGE	Comparison	SAGE	Comparison
Language arts				
Mean pretest	509.01	518.41	522.75	520.05
Mean posttest	563.32	558.72	572.80	558.32
Reading				
Mean pretest	515.62	519.53	523.27	528.32
Mean posttest	568.33	563.34	573.82	554.11
Mathematics				
Mean pretest	468.89	474.54	472.95	478.72
Mean posttest	524.23	517.03	522.01	506.22
Total				
Mean pretest	497.92	504.81	504.57	506.22
Mean posttest	552.34	546.82	556.72	539.73

TABLE 11
African-American and White Achievement in SAGE and Comparison Schools: Total Scale Scores for First-Grade Students

	1996–97		1997–98	
	Pretest	Posttest	Pretest	Posttest
SAGE schools				
African American	498.32	552.11	502.79	556.72
White	552.14	579.07	531.38	579.94
Comparison schools				
African American	505.40	547.44	510.07	539.73
White	531.26	579.23	528.60	569.02

reported Level 1 effects are the weighted averages of the within-classroom effects across the three models. The Level 1 results indicate that lower SES is related to lower posttest scores, and higher pretest scores are related to higher posttest scores (all statistically significant). When classroom pupil-teacher ratio is entered alone at Level 2, an increase of one in pupils per teacher can be expected to produce a 0.29 to 1.17 loss in average posttest performance. The results for all scores show this effect to be statistically significant.

Including both pupil-teacher ratio and SAGE participation in the analysis isolates the effects that SAGE might have beyond those produced by lower pupil-teacher ratio. The results show that once pupil-teacher ratio has been accounted for, SAGE has no significant effect on achievement. This suggests that the other SAGE interventions (i.e., rigorous curriculum, lighted schoolhouse, and staff development) do not have a significant impact on achievement in SAGE classrooms.

The third model, which combines class SES and class size, indicates that class SES has a significant effect on average posttest performance. The effect of a 1-point class average gain in SES equates to between a 10-point and 21-point gain on average posttest score, depending on the subtest. SES was measured on a 3-point family income scale; thus, a 1-point difference on average would be quite pronounced. The coefficients for pupil-teacher ratio remained similar to the first model for the 1996–97 scores but declined for the 1997–98 group; however, all remained statistically significant.

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

TABLE 12
HLM Results for First-Grade Students

	1996–1997				1997–1998			
Source	Total	Reading	Language arts	Math	Total	Reading	Language arts	Math
Level 1[a]								
Pretest	0.87	0.63	0.63	0.71	0.59	0.41	0.40	0.68
SES	–0.78	–3.73	–1.61	–3.20	–3.22	–4.44	–5.65	–2.44
Attendance	–0.21	–0.30	–0.28	–0.10				
Level 2								
A. Class size	–0.88*	–0.54*	–0.72*	–1.17*	–0.83*	–0.29*	–0.90*	–1.12*
B. Class size	–0.80*	–0.47	–0.63*	–1.12*	–0.65	–0.73	–0.64	–0.72
SAGE	1.28	1.79	3.57	1.54	2.99	0.20	4.23	6.41
C. Class SES	–18.08*	–15.99*	–16.13*	–21.03*	–12.96*	–10.41*	–12.97*	13.39*
Class size	–0.83*	–0.48*	–0.68*	–1.16*	–0.60*	–0.57*	–0.70*	–0.88*

*$p < .05$.
[a] Averaged within-class coefficients.

were collected only for SAGE first-grade classrooms. As reported earlier, SAGE students were organized into different types of classrooms: 15:1 regular classrooms, 15:1 shared-space classrooms, 30:2 two-teacher team classrooms, and 30:2 floating teacher classrooms. Data were available for 59 of the 15:1 classrooms and 31 of the 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. Therefore, the two 15:1 classroom types and the two 30:2 classroom types were merged for the purpose of analysis. This analysis, which used HLM, showed no statistically significant relationship between type of classroom organization and achievement on any of the subtests or the total score.

Limitations

When considering the results of the 1997–98 SAGE evaluation, several factors should be kept in mind. For example, the number of schools in the comparison group pool was reduced from 17 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 was addressed for the 1998–99 academic year by the addition of three comparison schools. Also, the test administered at the end of the 1996–97 school year to first graders showed significant ceiling effects, particularly for SAGE students. This was corrected in the second year. Thus, the 1996–97 first-grade effects are biased downward.

Effects on Classroom Environment and Teaching

The main thesis of the SAGE program is that reduced class size can increase student academic achievement. However, reduced class size cannot influence academic achievement directly. Reduced class size is mediated by classroom events. It must influence what teachers and students do in the classroom before it can possibly have any effect on students' learning. To fully understand achievement effects in relation to reduced class size, it is necessary, therefore, to examine classroom changes brought about as a result of reduced class size. The relationship of classroom events to reduced class size, the principal SAGE variable, was examined through data obtained from teacher interviews, classroom observations, teacher logs, and teacher questionnaires. The results reported here are derived from the teacher questionnaires and the teacher interviews, because they most directly focused on classroom changes related to reduced class size. The results are from both first- and second-grade teachers in SAGE classrooms.

Teacher Questionnaires

Teacher questionnaires were completed by SAGE teachers during spring 1998. The questionnaire, which elicits perceptions regarding classroom teaching, mathematics, reading, language arts, family involvement, professional development, and overall SAGE satisfaction, was returned, in usable condition, by 150 first- and second-grade SAGE teachers (first grade: 72 teachers; second grade: 78 teachers).

The classroom teaching section, which contributes to this examination of classroom events in reduced class size, consists of 11 items. The teachers were to indicate the level of their agreement with each item and then select and rank the 3 that represented the most significant ways their teaching had been affected by reduced-size class. The results of these two analyses for the total group of first- and second-grade teachers and for each type of SAGE classroom follow.

Total questionnaire results. Table 13 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 were 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, were 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. The most important classroom product of reduced class size, according to these data, is individualization. Class participation, which can be viewed as a type of individualization because individual students receive 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 suggest that the type of teaching used in small-sized classes is teacher-centered, teacher-controlled teaching.

The teacher behaviors receiving the lowest ratings and rankings again suggest that teachers are

TABLE 13
Total Questionnaire Results (N = *150)*

Item	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Ranking percentage
1. More time teaching	1	1	7	42	49	17
2. Covered more content	0	1	5	40	54	8
3. Integrated content more often	0	0	12	50	39	4
4. Covered content in more depth	0	1	9	46	44	5
5. More time individualizing	0	0	3	30	68	25
6. More discussing and sharing	0	0	2	42	56	11
7. More hands-on activities	0	0	7	56	38	10
8. More use of prior knowledge	0	1	15	49	36	4
9. More problem solving and creating	1	1	6	51	42	6
10. More cooperative groups	0	4	19	42	36	5
11. More choice in activities	0	4	14	47	35	5

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 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 classroom 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 with the other teachers, the 15:1 shared-space teachers used hands-on activities less, the 15:1 regular teachers used problem solving more, those in 30:2 two-teacher teams used students' prior knowledge more, and those in 30:2 floating teacher classrooms used more time for instruction.

Teacher Interviews

Twenty-eight SAGE teachers were interviewed, either individually or in teams, in spring of 1998. They were chosen to represent different geographic areas of the state, grade levels, and types of SAGE classrooms. 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 temporary wall), 2 teachers; 30:2 two-teacher team (two teachers teaching 30 students), 10 teachers; and 45:3 three-teacher team (three teachers teaching 45 students), 6 teachers. The interviews, which lasted from 20 minutes to more than 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 3 of SAGE. Results regarding these three areas follow.

Teaching change. Each of the interviewed teachers indicated that his or her teaching had changed as a result of having a small-sized class. The areas mentioned most frequently were knowledge of students, discipline, instruction, individualization, and learning activities. These findings reinforce and extend the teacher questionnaire results.

1. *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.

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

3. *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 "paperwork" because of having fewer students also resulted in more time available for teaching.

4. *Individualization.* All interviewed teachers 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 the 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. Furthermore, 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.

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 the reduced class size, results in reduced need for remediation of instruction later, they believe.

5. *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 more hands-on manipulative activities, more enrichment-type activities, more interest centers, and more cooperative groups than they had previously used. They used more activities that 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 in which 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.

Student learning. All of the interviewed teachers stated that their students' achievement has increased considerably as a result of being in a small-sized class. They reported 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 reported that they were able to expand and deepen students' learning. They were 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 were 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 learned at the most rapid rate were those who had been in SAGE classrooms the previous year. The teachers said that these students were instantly recognizable as soon as school began in the fall.

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 identify general problems, such as the need for teacher in-service and more hands-on activities, 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 paperwork, 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.

Classroom Events Summary

The results from the 1997–98 teacher questionnaires and teacher interviews 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 appears to be a focus on students as individuals. Many, if not most, of the techniques and methods that teachers use may be the same techniques and methods that they have used in normal-sized classrooms. The difference is that now the techniques and methods are directed at individuals much more frequently. Teachers know each student's learning needs, they correct misunderstandings instantly, and they move ahead when the time is right. This attention to individuals is implemented 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.

Although all conclusions about teaching small-sized classes must be tentative at this time, a model of teaching such classes is beginning to emerge. The model contains many elements, but it emphasizes individualization. It speculates that having fewer students permits teachers to know students better, results in more time for instruction (because it reduces misbehavior and time needed to manage the classroom), and leads to greater teacher satisfaction and pleasure regarding teaching. These three elements—student knowledge, instructional time, and teacher satisfaction—come together to permit more individualized instruction. Greater teacher satisfaction and more instructional time, both of which are products, in part, of less misbehavior (which in turn is related to greater knowledge of students), permit somewhat more use of hands-on activities involving regular teacher-centered methods. The outcome of this heavy focus on individuals is more content coverage and, the model theorizes, more student achievement.

The model applies across the school year, grade levels, and types of SAGE classrooms. Few differences over the year, in grade level, and in type of classroom were revealed in classroom data from any of the instruments. 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. Furthermore, one might expect differences between fall and spring data, particularly for teachers new to the SAGE program. And differences might be expected among types of SAGE classrooms because having responsibility for 15 students (as occurs in 15:1 student-teacher ratio classrooms) is clearly different from two teachers sharing responsibility for 30 students (as occurs in team-taught classrooms). While these three areas will be investigated further as the SAGE evaluation continues, a tentative explanation is that the apparent lack of major differences in teaching small classes over time, by grade level, and by types of SAGE classrooms may be related to the possibility that changes in teaching occur quickly at the start of the school year, and they occur for most teachers in a similar way. The teacher-centered approach that they had been using is repackaged for individual distribution. Small class size does not demand learning something new; it permits teachers to do what they know is the right thing to do.

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. Case studies of selected SAGE schools and classrooms are now being conducted to provide a clearer picture of life in small-sized classrooms.

Discussion

The impact of reduced class size in Wisconsin's SAGE program appears to be generally consistent

with the results reported by the Tennessee STAR study. The 1997–98 first-grade SAGE students gained about 0.2 standard deviations for class size reductions that were slightly larger than the Tennessee reductions. Students in 1997–98 SAGE first-grade classrooms scored significantly higher in all tested areas. The results reported for 1996–97 showed somewhat smaller effects, but the results were biased downward by ceiling effects on the end of year test instrument.

Consistent with the STAR findings as well, the achievement benefits of SAGE's small class sizes are stronger for African-American students. Indeed, 1997–98 SAGE first-grade posttest results reveal that African-American SAGE first-grade students closed the "achievement gap" with White students that was revealed in the first-grade pretest results. If these findings are confirmed in subsequent evaluations, a strong case could be made for reducing class size in the early grades as an effective means of addressing the achievement gap between African-American and White students, one of the most persistent and troubling problems in American public education.

It is also worth noting that an analysis of 15:1 and 30:2 first grade classroom achievement results did not reveal statistically significant differences on any of the subtests or total score. If this finding holds up in future analyses, it suggests that the benefits of reducing class size may be achievable without the attendant capital costs of building additional classrooms.

The results of analyses of classroom environment and teaching practice data are also of considerable interest. These analyses suggest that teachers in SAGE classrooms have greater knowledge of each of their students, spend less time managing their classes, have more time for instruction, and individualize instruction using a primarily teacher-centered approach.

Notes

For further information on the SAGE project, see the following Web site: **http://www.uwm.edu\soe\centersprojects\sage**.

References

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

Comprehensive Test of Basic Skills (4th ed.). (1991). New York: CTB Macmillan/McGraw-Hill.

Maier, P., Molnar, A., Smith, P., & Zahorik, J. (1997). *First year results of the Student Achievement Guarantee in Education program*. Milwaukee, WI: Center for Urban Initiatives and Research.

Molnar, A., Smith, P., & Zahorik, J. (1998). *1997–98 results of the Student Achievement Guarantee in Education (SAGE) program evaluation*. Milwaukee, WI: School of Education.

Authors

ALEX MOLNAR is a professor at the Department of Curriculum and Instruction, University of Wisconsin–Milwaukee, Enderis Hall, Third Floor, 2400 East Hartford Avenue, Milwaukee, WI 53211. He specializes in urban education, education policy, and curriculum and instruction.

PHILIP SMITH is a professor of educational psychology at the School of Education, University of Wisconsin–Milwaukee, Enderis Hall 777, 2400 East Hartford Avenue, Milwaukee, WI 53211. He specializes in applied measurement and research methods.

JOHN ZAHORIK is a professor at the Department of Curriculum and Instruction, University of Wisconsin–Milwaukee, P.O. Box 413, Milwaukee, WI 53201. He specializes in instructional research and theory.

AMANDA PALMER is a doctoral student at University of Wisconsin–Milwaukee, Enderis Hall 366, 2400 East Hartford Avenue, Milwaukee, WI 53211. She specializes in urban education, educational psychology, and research and evaluation.

ANKE HALBACH is a doctoral student at University of Wisconsin–Milwaukee, Enderis Hall 366, 2400 East Hartford Avenue, Milwaukee, WI 53211. She specializes in urban education and curriculum and instruction.

KAREN EHRLE is a doctoral student at University of Wisconsin–Milwaukee, Enderis Hall 366, 2400 East Hartford Avenue, Milwaukee, WI 53211. She specializes in urban education and curriculum and instruction.

Manuscript received January 15, 1999
Revision received February 23, 1999
Accepted February 24, 1999